Communication and Group Behavior in Economic Decision Making

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften
(Dr. rer. pol.)

durch die Fakultät für Wirtschaftswissenschaften der
Universität Duisburg-Essen
Campus Essen

vorgelegt von
Name: Klemens Keldenich
Geburtsort: Köln
Essen 2013

Tag der mündlichen Prüfung:	16.07.2013
Erstgutachterin:	Prof. Dr. Jeannette Brosig-Koch
Zweitgutachter:	Prof. Dr. Erwin Amann

Acknowledgments

First of all, I would like to thank my supervisor Jeannette Brosig-Koch for the continuous support of my Ph.D study and research. Her patience and immense knowledge greatly improved this thesis, and her openness towards new ideas and innovative experimental designs was very important for me.

In addition, I would like to thank my second supervisor Erwin Amann for interesting and valuable discussions during my years at the University of Duisburg-Essen. Werner Nienhüser kindly completed the examination committee.

I am also grateful to the Ruhr Graduate School in Economics, not only for providing financial support but mainly for creating an atmosphere where learning about economics comes naturally. Special thanks go to the members of my cohort: Claudia Burgard, Frauke Dobnik, Regina Flake, Maarten van Kampen, Jonas Keil, and Marcus Klemm. The same thanks go to the colleagues at the University of Duisburg-Essen: Timo Heinrich, Christoph Helbach, Johanna Kokot, Nadja Schwarz, and Guanzhong Yang. It was a pleasure working with all of you!

As the papers in this thesis are very much a collaborative effort, my various co-authors deserve special mention: I have greatly benefitted from each of you in different ways.

Finally, I would like to express my gratitude to my mother Monika, my sister Katrin, and my fiancée Eva. Your support was and is the bedrock for my work. Thank you all very much.

Contents

Introduction

Communication and group behavior

The well-known quote “One cannot not communicate” (Watzlawick et al., 1967) highlights the importance of communication as a central element of human interaction in general and therefore also of many economic transactions. Communication as the exchange of information between two or more individuals is inextricably linked with group decision making. This thesis aims to investigate the connection between both communication and group behavior on one side and economic decision making on the other side from several angles. For this investigation, five empirical studies are conducted, utilizing methods of experimental economics: four of the five studies presented use a laboratory experiment, while the remaining study takes advantage of data gathered from a "natural experiment", namely a TV quiz show.

In general, communication can be analyzed from two starting points: The first one takes communication as the independent variable, i.e. a cause of changes in a certain outcome. Examples for such outcomes might be individual behavior, institutions, or allocations. The second starting point turns this around and looks at communication as the dependent variable, i.e. how communication itself is affected by certain institutions or behavior. Both approaches are interesting from an economic point of view and examples for them will be presented in the thesis at hand.

Communication as the independent variable

Of course, communication is not only interesting on its own but usually serves as a transmission channel for new information. This makes it sometimes difficult to investigate communication effects per se, as possible changes in behavior caused by communication cannot be separated from effects on behavior caused by the new information. Still, looking at communication in these situations is a useful way to open up the “black box” of decision making and understand how relevant information is transferred between agents. Chapter 2 of this thesis belongs to this area: groups of two or three people have to make one binding decision about a risky choice and therefore need to exchange their preferences via direct face-to-face communication. Here, the exchanged information is new to the contestants and will influence the final decision. In addition to that, parts of chapter 5 can also be classified as looking at this kind of information. In the study presented there, the effect of increasing the amount of information agents receive about their consumption decisions is analyzed. Again, this information is new to the decision makers, and changes their information situation.

A special kind of communication from an economic point of view is so-called “cheap talk”: This type of communication is cheap in the sense that it does not have direct payoff implications and is an active topic of research in economic decision making (see for example Crawford, 1998 for a survey). Chapter 1 can be sorted into this range: three-person groups can communicate with each other, but only one pre-determined group leader makes the binding decision for all. Thus, the information exchanged is not payoff-relevant.

A second way how communication without payoff-relevant information might influence decisions is through framing. Framing denotes the presentation of information without influencing the relevant pieces of data. This phenomenon has been an active research topic in the economic literature for some time, see for example Kahneman (2003) for an integration of this concept into the general field of behavioral economics. Chapter 5’s study contains elements of this kind of information, as it looks at the effect of different frames on consumption decisions. However, in the decision situation analyzed in this study, framing has no influence on behavior.

The interplay of group decision making and communication is also situated in the area of communication as the influencing variable. As the communication usually takes place in a group, the effect of these two influences can oftentimes not be distinguished. Chapter 1 aims to take a first step at disentangling group membership and communication by manipulating the group’s decision rule.

Communication as the dependent variable

When looking at communication as the outcome variable, several potential influences immediately come to mind: communication medium (see for example Brosig et al., 2003 for the analysis of the influence of different communication media on cooperation), possible time constraints, different kinds of participants, etc. might influence the communication. From an economic point of view, the financial incentives aligned with the communication have a special place for this question. Chapter 3 looks at communication as teaching: the one-directional exchange of information from one agent with high information to one with low information. The effects of monetary incentives on this kind of communication are analyzed, with obvious implications for the question of how to improve the knowledge transmission taking place. In chapter 4, a different kind of influence on communication is looked at. In a group, different bilateral connections are allowed and the resulting network structure affects how fast and efficient information is exchanged in the group.

Of course, the analyses presented in this thesis cannot exhaustively catch the importance of communication and group behavior on economic decision making. The framework presented in this introduction is supposed to enable a better orientation where the specific research is situated in the bigger picture. The thesis proceeds as follows: this chapter will continue with a short overview of the method of laboratory experiments in economics and then conclude with a summary of the studies comprising the thesis. Chapters 1 to 5 will present the five studies themselves, while the last chapter offers a short conclusion.

The experimental method in economics

The use of laboratory experiments as a standard tool in economics is a relatively recent development. Croson and Gächter (2010) e.g. state that "Some years ago, it was still common to argue that economics could never be an experimental science but was confined to be purely observational or theoretical." However, the methodological advantages offered by laboratory experiments are increasingly acknowledged and used in economic research. Compared to other empirical research methods, the main advantage of doing laboratory experiments is the ability to control the conditions under which the decision of interest takes place. This allows the experimenter to keep the background conditions constant while only changing the one factor (the "treatment") which is of interest[1].

Before the widespread use of experiments in economics, there were nevertheless some early strands of experimental research in economics. Roth (1993) names three different broad areas where early experiments were conducted: The testing of theories of individual choice (starting with Thurstone, 1931), the testing of hypotheses from game theory (starting with Flood, 1958), and experimental work in Industrial Organization (with an early contribution by Siegel and Fouraker, 1960). While these areas have developed over time and are still important in experimental economics, almost all other areas of empirical inquiry in economics now also use experimental methods. Samuelson (2005) even suggests that "experimental economics is currently making its transition from topic to tool", likening it to game theory or econometrics, which made similar transitions.

The method of laboratory experiments is of course well established in other academic disciplines. Although the social sciences generally do not use experiments as a standard method, an exception is the research done in social psychology. This work is oftentimes closely related to the work done in experimental economics. However, while many methodological similarities exist, some major differences are also worth pointing out. This is especially important when comparing results from both disciplines which deal with similar research questions. The most important differences regard the payment of the subjects and the question whether to deceive subjects (see for example Ortmann and Hertwig, 2002 for a discussion). As to the first point, economists usually use monetary incentives to motivate subjects, claiming that this way the desired preferences can be induced. Social psychology, however, usually abstains from monetary rewards (oftentimes using course credit which is given regardless of the decisions in the experiment). The second difference concerns the degree of transparency towards the participants in the experiment. While economists maintain that subjects should never be actively deceived, this is common practice in social psychology. The reasons for the refusal to mislead subjects stems (mainly) from the possible subject pool effects. If the subjects later on realize that they have been deceived (as they usually do in psychological experiments in the debriefing), this information might spread to the whole subject pool and thus change the expectations future subjects have

[1] Of course, perfect control - i.e. an exact replication - is not possible. The experiment has to be done either with different participants or - when the same participants are used - their experiences have necessarily changed.

in the experiment. If a subject enters the laboratory thinking that he will be deceived, it is very hard for the experimenter to understand or control the motives behind subjects' behavior. Psychologists, however, maintain that certain research questions can only be answered using deception and that it is sufficient to inform the subjects after the experiment.

Today, experiments are used as a method to answer a wide variety of economics questions. Following Roth (1995), the use of experiments in economics can be broadly classified into three areas: Testing theories (called by Roth "speaking to theorists"), uncovering behavioral regularities not described by existing theories ("searching for facts"), and delivering and improving policy advice ("whispering in the ears of princes")[2]. For each of these areas, prominent examples can easily be found: Regarding the first two areas, one can look at the emergence of theories of other-regarding behavior which now complement or even replace those where an agent's utility is only influenced by his own payoff. These theories (for examples see Fehr and Schmidt, 1999, Rabin, 1993, Bolton and Ockenfels, 2000) originated from observed behavioral regularities which violated the assumption of perfect selfishness. Thus, while one theory was tested (and found lacking under certain circumstances), a new one was developed to better organize the experimental data.

Regarding the third area - giving policy advice - a prominent example can be found in the area of market design. Many real world markets have been the subject of experimental analysis, including the market for US physicians, school choice mechanisms (see Roth, 2002 for concrete examples), or online auctions (Ockenfels and Roth, 2006). Of course these examples do not mean that experimental methods replace more traditional empirical methods, but rather complement them.

Despite this widespread use and acceptance of experimental methods in economics, there are still controversies regarding the utility and scope of the approach. Levitt and List (2007), Levitt and List (2006), and Schram (2005) are three examples that describe several problems of (laboratory) experimental economics which primarily concern external validity, i.e. the transferability of results to a different context. However, Croson and Gächter (2010) pick up these criticisms to argue that laboratory experiments can either be used to address most of them or that their methodological concerns are valid for all empirical analyses.

[2]Roth (2010) revisits these three areas and gives a short overview on the accomplishments of experimental economics in each of the three.

For the purpose of the studies presented here, the experimental method is especially well suited. The laboratory environment allows a manipulation and control of all communication between the relevant agents. This is very difficult to achieve in field situations, regardless of the type of communication analyzed. As the communication can be recorded (both written as in chapter 1 and video-stream as in chapter 3), a later in-depth analysis is possible and can yield valuable insights.

The five studies

This thesis analyzes several aspects of group decisions and communication in economic decision making. **Chapter 1** deals with communication as "cheap talk"; i.e. communication that is not relevant to the decision maker's monetary payoff. Instead, it allows the exchange of preferences between the agent making the decision and other agents who are influenced by her decision. In this way, it changes the decision maker's perspective of belonging to a group and influences results. Specifically, chapter 1 aims to isolate the effect group membership alone has on decision making. The research uses a modified dictator game as a vehicle and employs a within-subject design to differentiate individual decision from group decision with and without communication in the group.

The results suggest that a possible group membership effect is influenced by the degree of group membership saliency. The within-subject design uses 4 stages: in stage 1, each subject decides individually; in stage 2, the subjects are divided into groups of three and one person is selected at random from each group to make the decision (the "hierarchical decision rule"). In stage 3, additional pre-play communication in the group is allowed before the decision and, in stage 4, the decisions are again made on an individual basis. Interestingly, the dictators behave more selfishly when group members are not allowed to communicate. However, if groups are allowed to communicate, decisions do not differ from individual choices. Chat content shows that groups are concerned with reaching a consensus, even though talk is "cheap" and only one group member will make the binding decision.

Chapter 2 also looks at the decision making process in groups. However, in contrast to the first study, no clear decision rule is determined a priori. This means communication in the group is not only "cheap talk": To reach a decision, the group members have to exchange their preferences and then find a way to aggregate them. The study uses data from the game show “Quiz Taxi”, where groups face the decision whether to bet the winnings they have acquired on a final “double or nothing” question. The decision is made by groups of two or three persons.

Although this study dos not use a laboratory experiment, the situation constitutes a "natural experiment" in the sense that rules and payoffs are clearly determined and the subjects' behavior can be monitored closely. This set-up enables the decision making process to be studied by observing group communication. There is a strong correlation between communication content and the final choice, indicating that, from the contestant's perspective, the decisions are rational and that the context is an important factor in the final decision. This is particularly so for individual valuations of the money at stake. More extensive discussions help to make the right decision. As contestants do not apply to go on the show, they represent a less selected sample than those in previous game show studies. Overall, the contestants show risk averse behavior, suggesting CRRA-parameters[3] larger than 1. The study also shows some heterogeneity in attitude to risk. Contestants who do better in the show are more likely to go for the risky option, because they are more knowledgeable and more confident. All-female groups are less likely and three-person groups more likely to choose the risky option.

While the first two chapters study group decisions and the communication with regards to the outcomes of these processes in the group, **chapter 3** looks at the communication itself. The focus there lies on communication as the transfer of knowledge. This knowledge transmission is an important part of business and private life, with the education sector being the foremost example. One important question in this field is how monetary incentives influence the process of knowledge transmission. This study uses a laboratory experiment to analyze the effectiveness of performance-based monetary incentives in the teaching process.

The process of knowledge transmission is recreated using a video-stream. Four different teacher payment schemes are compared: A fixed wage, a piece rate for student success, a bonus if the student reaches a certain threshold, and a tournament comparing three teachers by means of their student's performance. Furthermore, the experiment is done with two different subject pools: prospective teachers and regular students. This serves to address the question of a possible intrinsic motivation for teaching by the prospective

[3]Constant relative risk aversion.

teachers. The resulting videos are analyzed to see directly what part of the teaching might change in reaction to the different incentive schemes.

To control for student ability, the share of transmitted knowledge from teacher to student is evaluated in addition to the absolute amount of student knowledge. The results indicate that prospective teachers do not react to monetary incentives. There is no difference in student performance across treatments. However, regular students in the teacher role do react in the expected way: Teachers transmit a significantly higher share of their knowledge when paid according to student performance. These results suggest that transferring results concerning incentive schemes to the educational sector might be problematic, as at least prospective teachers might react differently to monetary incentives.

In a similar vein to this study, **chapter 4** also tries to illuminate what influences the process of communication in a group. In this study, however, monetary incentives are kept constant for all group members and are not varied across treatments. Instead, a laboratory experiment is used to systematically vary the communication structure - i.e. the possibility to communicate with another group member - in the group. Four of these network structures are analyzed in their effects on knowledge dissemination. The effect of exogenous network structures, which are modeled as five-actor groups, in a non-strategic situation is investigated. Every individual represents a node and possesses some private information. The different network structures are characterized by a different number and variance of links between the individual nodes. The experiment addresses some characteristics of typical situations that arise in the public promotion of R&D cluster and network formation: An initial network structure has developed over a long time span and policy measures result in a change in the structure of links between the actors. These policy measures influence the effectiveness of the information flow in the network structures in a way that is not clear from the beginning.

The experimental results suggest that the different network structures do indeed influence the way information is exchanged. Both too many possible links (causing a coordination problem) and too few possible links (introducing bottlenecks) are harmful. The participants in all network structures learn over time and achieve a faster exchange of information in the later rounds. Furthermore, participants' behavior seems to converge to equilibrium in later rounds, as the same connections are formed over and over again. These results suggest that when influencing communication structures, one has to be careful to balance the positive and negative effects of adding more communication possibilities.

Finally, **chapter 5** looks at communication as a means of providing information to a decision maker and analyzes how different kinds of information are used. Deviating in some way to the studies presented so far, this analysis mainly looks at individual decision making and only lightly touches on the possible influence of communication with other actors.

A laboratory experiment is used to analyze the impact of different types of information on consumption and savings behavior. Based on a buffer stock savings model, three treatment dimensions are used: The amount of information subjects receive about the likelihood of income shocks, whether subjects are informed about other people's beliefs about these shocks, and the framing of shocks. The results reveal that - even with little information about the random element determining the income shock - consumption decisions are surprisingly close to the optimal consumption path. If at all, more information rather worsens than improves consumption behavior. Nevertheless, in line with the theoretical prediction, observed behavior is robust to the framing and other people's beliefs about income shocks. Given that actual decisions are significantly correlated with the optimal consumption amount (and not with easier accessible variables like cash-on-hand) suggests that subjects do not simply use naive heuristics to determine their consumption.

Chapter 1

Group membership and communication in modified dictator games [4]

1.1 Introduction

Many important economic decisions are made by teams rather than by individuals. Examples for these decisions include decisions about consumption and savings, "virtually all significant strategic decisions by corporations" (Cooper and Kagel, 2005), central bank decisions regarding monetary policy (Blinder, 2007), or investment decisions by mutual funds (Prather and Middleton, 2002). To capture all relevant aspects of decision situations like these, it is therefore important to see if groups behave differently than individuals. Recently, group behavior has become the focus of many economic studies which can be divided into two main categories. The first one looks at groups which have to come to a consensus decision and where no payoff conflict is present. Following Sutter (2009), this approach is called team decision making. The second, more recent area of investigation examines whether group membership alone is sufficient to cause a change in behavior.

The results of team decision making studies usually show that teams are closer to the standard game-theoretic predictions in the ultimatum game, thus choosing lower offers as proposers and accepting lower offers as responders (Bornstein and Yaniv, 1998), and send

[4]See Keldenich, K. (2012): Group membership and communication in modified dictator games, *Ruhr Economic Papers 322*.

and return smaller amounts in an investment game (Cox, 2002 and Kugler et al., 2007). Furthermore, teams exit a centipede game earlier than individuals (Bornstein et al., 2004) and show higher levels of rationality in beauty contests, thus outperforming individuals (Kocher and Sutter, 2008). While most of these studies show that groups are more rational and / or selfish, Cason and Mui (1997) find that groups are more generous in the dictator game.[5] Similar results can be found in studies of non-strategic tasks. For example, a study on portfolio selection by Rockenbach et al. (2007) shows that teams are better at making the trade-off between risk and higher expected payoff and a study by Charness et al. (2007a) demonstrates that teams are better than individuals at following the principles of Bayesian updating.

The second area of study looks at the effect of group membership alone. Here, one can differentiate between naturally occurring groups and induced group membership. As an example, Bernhard et al. (2006) study different native social groups in Papua New Guinea using a dictator game with third-party punishment. They find ingroup favoritism in sharing decisions and norm enforcement. From a study of different ethnic groups in Vietnamese village communities, Tanaka et al. (2006) suggest that the effect of group membership depends on the respective status of the groups concerned. Finally, Goette et al. (2006) look at different platoons in the Swiss army as groups and find that cooperation in a prisoner's dilemma game increases with ingroup players. Studies with induced group membership allow more control by the experimenter, thus making it possible to vary the strength of the group membership. Eckel and Grossman (2005) use a public good game to this end and find that "just being identified with a team is, alone, insufficient to overcome self-interest." However, strengthening group identification, for example through problem solving exercises, leads to higher contributions and less free-riding in the public good game. Similar results are found by Charness et al. (2007b) using prisoner's dilemma and battle-of-the-sexes games. When group membership is sufficiently salient, it significantly alters individual behavior.

This study lies at the intersection of these two literature areas. While it uses communication and payoff commonality as in team decision making studies, it isolates the group membership effect by introducing a hierarchical decision rule. This is innovative in two respects: firstly, the decision rule allows a clear comparison between decisions made by an individual as part of a group and acting alone. Secondly, the content of the communication

[5]Although a follow-up study by Luhan et al. (2009) arrives at the opposite result.

is recorded, so that possible reasons for the effect of group membership can be analyzed. The paper mainly focuses on the difference in the way subjects behave when they are part of a group compared to when they are acting alone. The main results of this paper are as follows: While introducing group membership through payoff commonality makes subjects more selfish, the decision in a group with pre-play chat communication is not different from the individual decision. Furthermore, when communication in the group is allowed, a consensus decision is actively sought after, even though the decision rule does not need the agreement of all group members. The communication content shows that the group members are aware of the decision rule, but still seek to influence the final decision.

1.2 Experimental design

1.2.1 Structure

To study the effect of group membership on behavior, a four-stage experiment with a within-subject design is conducted. In each stage a modified dictator game (introduced by Andreoni and Miller, 2002) is used as the base game. Individual behavior in the modified dictator game has been analyzed by Bardsley (2008) and List (2007), who both find that modifying the game's structure, i.e. the addition of a taking option, influences outcomes significantly compared to the standard dictator game. The modified dictator game is used here as it allows a wider range of behavior for the subjects, making it easier to detect the possible effect of group membership. At the end of the experiment, one stage is randomly chosen and payed out[6]. Upon arrival, subjects are randomly divided into Dictators and Receivers and retain these roles throughout the whole experiment. The roles are called "Type A" (Dictator) and "Type B" (Receiver) for the participants. Before each stage, new instructions about the stage are distributed. All participants receive the same instructions[7].

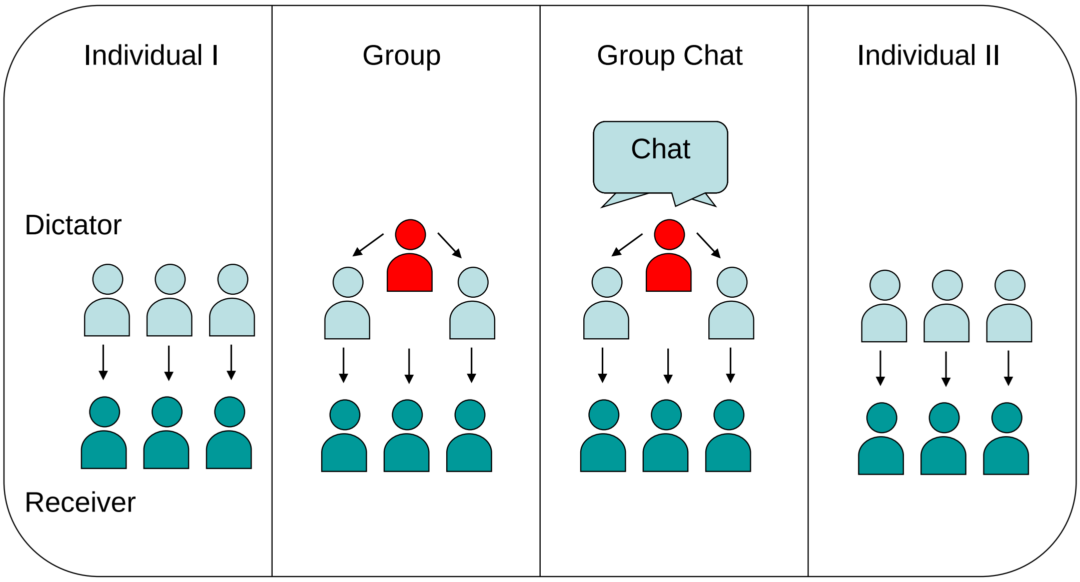
In the first stage (called *Individual I*, see Figure 1.1 for an overview), both Dictators and Receivers are given EUR 12 as an endowment. The Dictators decide if they want to keep the endowment, take away money from their randomly assigned Receiver, or give him money from their own endowment. Transfers are only allowed in integers of 1 Euro. Thus,

[6] Baltussen et al. (2011) provide an overview of the commonly used method of paying only one randomly selected decision from a sequence of decisions.

[7] See Appendix for a complete set of instructions.

possible divisions range from EUR 24 for the Dictator and EUR 0 for the Receiver to EUR 0 for the Dictator and EUR 24 for the Receiver.

Figure 1.1: Overview experimental design



In the second stage (*Group*), the Dictators are randomly divided into groups of three and each Dictator is matched anew to a Receiver. Every participant receives a new endowment of EUR 12. Now, Dictators have to decide for the whole group including themselves how to split the money between each Dictator and his assigned Receiver. A variant of the strategy method[8] is used to determine the binding decision for all group members: Each Dictator in the group has to decide on the split and one of the decisions is randomly chosen afterwards to be the binding one for the group. This decision is then applied to all members of the particular group. Dictators do not get any information on which decision was binding in the group. The matching of one Dictator to one Receiver is kept to avoid that Dictators change their behavior because they are now dealing with a group instead of dealing with an individual. Sutter et al. (2007) for example argue that when interacting with a group instead of an individual, an out-group scheme is recalled which renders the interaction "competitive, deceitful and aggressive" (see also Pemberton et al., 1996).

In the third stage (*Group Chat*), the Dictators are again randomly divided into groups of three, where the subjects are labeled "Number 1", "Number 2" and "Number 3". Each

[8]See Selten (1967).

group can communicate internally via electronic chat for five minutes. The chat is set up in a way that all group members can see all messages, it is not possible to write a message to a single group member. All messages stay visible during the five minutes, so one group member can look back to the beginning of the chat and see what has been written. After the chat, the subject labeled "Number 3" makes one decision which is binding for all group members, including himself. This decision rule is known to all group members before the chat. In addition, the group members labeled "Number 1" and "Number 2" are asked how they would have decided in the position of "Number 3". An electronic chat is used as it is easy to record for subsequent analysis, retains a high level of anonymity (compared to face-to-face or audio chat), and is natural for participants. The subjects were informed that they can communicate only just prior to the third stage, they did not yet know about the future communication in the second stage.

The fourth and final stage (*Individual II*) consists again of an individual modified dictator game with new random matching. After the last stage, Dictators and Receivers are informed about the result of every stage and one stage is randomly chosen to be payed out.

1.2.2 Procedures

The experiment was programmed with z-Tree (Fischbacher, 2007) and conducted at the "Essen laboratory for experimental economics" (elfe) in June 2010. Subjects were recruited by the online recruiting system ORSEE (Greiner, 2004). After the experiment, payoffs were made separately. The participants were undergraduate students from the University of Duisburg-Essen; their fields of study included business administration, economics, law, languages, and others. The experiment as a whole lasted about 60 minutes, including payoff time. Average payoff was EUR 12 with the highest payoff being EUR 24 and the lowest one EUR 0. Four sessions with 24 participants each were conducted, leading to a total of 96 subjects.

1.3 Research questions and expectations

1.3.1 Theoretical background

Standard economic theory focuses on individual-level incentives in decision making and thus has no place for group membership effects. If in addition the assumption of selfishness holds, the Dictators should take all of the endowment in all four stages, leading to the following hypothesis:

Selfish: *"The Dictators will take all of the endowment in all 4 stages."*

Of course, prior research using the modified dictator game (see Bardsley, 2008 and List, 2007) has shown that such purely selfish behavior is seldomly observed. Still, it is useful as a simple and clear baseline to measure behavior against. While social preferences introduced other people's payoffs into an individual utility function and therefore leave the selfish assumption, this is not affected by group membership either. So the existence of some kind of social preferences can be used to explain why Dictators do not take the maximum amount for themselves, but is not sufficient to account for changes between the stages. If therefore group membership has no influence on behavior, the following hypothesis can be formulated:

Group: *"The group leaders' decisions will not change between the stages* Individual I *and* Group.*"*

In contrast to standard economic theory, social psychology has a long tradition in analyzing group membership effects. The "Group Polarization Hypothesis" based on two underlying theories is a possible explanation for these effects. This hypothesis, first presented by Moscovici and Zavalloni (1969), states that "the average postgroup response will tend to be more extreme in the same direction as the average of the pregroup responses." It is stated for cases in which communication among the group is allowed.

This shift may have an informational or a normative explanation. The dominant informational explanation is the "Persuasive Argument Theory", while the normative explanations stress the tendency of group members to compare themselves with others and the wish to be perceived positively and is formalized in the "Social Comparison Theory" (SCT). According to the "Persuasive Argument Theory" (PAT), people are influenced by the number and persuasiveness of pro and contra arguments that they can recall from memory when making decisions. In a group, arguments are pooled, so the initial positions are enhanced by more arguments. This influence consists of the observation that a discussion generates arguments which predominantly favor the initially preferred alternative. The "Social Comparison Theory" provides a second theoretical explanation for the Group Polarization phenomenon. It states that people are motivated both to perceive and to present themselves in a socially desirable way. Furthermore, people tend to perceive themselves as more favorable than what they believe to be the average tendency. According to the "Group Polarization Hypothesis", group discussion moves the decision into the direction of initial tendency. This means that the group's leader is influenced by the other group members' preferences which they can communicate during the chat. The amount of money divided is held constant on an individual level, so there is no incentive to change behavior. In addition, no consensus decision in the group is needed, as one group member is randomly picked to make the binding decision. Thus, no compromise is necessary and every Dictator can decide purely according to her preferences. This ensures that the group membership is the only variable which changed. Following this, the third hypothesis is defined as follows:

GroupChatA: *"The group leaders' decisions will change between the stages* Individual I *and* Group Chat *in the direction of the median decisions of all group members."*

The same should happen to the hypothetical offers by the group members who do not have the leader role and whose offers are therefore not payoff-relevant. These hypothetical decision may be even more influenced by the wish to appear socially desirable, because this does not incur any costs for the group members: As their decisions do not have payoff consequences, they do not have to suffer the utility loss associated with receiving less money. They can conform to the social norm without costs to themselves. From this reasoning, another hypothesis can be derived regarding the behavior after the pre-play communication:

GroupChatB: *"The group members' hypothetical decisions will change between the stages* Individual I *and* Group Chat *in the direction of the median decisions of all group members. This difference will be more pronounced than the one for the group leaders."*

In stage *Individual II*, the subjects decide individually again. Looking at the behavior in this stage, it is necessary to distinguish between the two possible underlying causes of the "Group Polarization Hypothesis". If one follows the "Social Comparison Theory", a possible change in behavior is only due to the desire to be perceived in a certain way by the group members. In stage *Individual II* of the experiment, this is no longer the case and the leader's offer should fall back to the initial offer he made in stage *Individual I*:

SCT: *"The Dictators' decisions will not change between the stages* Individual I *and* Individual II.*"*

However, if the "Persuasive Argument Theory" is the cause of the polarization, the arguments presented by the other group members have influenced the leader. Assuming that the subjects can still recall the arguments from stage Group Chat, the change in behavior should be permanent, meaning that there will be a change in the decisions made between the two individual stages. Additionally, as the group members have been influenced by the very same arguments, the final hypothesis regarding Dictator behavior is defined as follows:

PAT: *"The Dictators' decisions change from stage* Individual I *to* Individual II *in the direction of the group decision."*

The effect of group membership on social preferences which incorporate other people's payoffs into a utility function has been analyzed by Chen and Li (2009) using several games[9]. In short, all of their results are compatible with the hypothesis that participants are more altruistic towards an ingroup match. The analysis presented here uses a similar theoretical framework of group membership or group identity, but differs with regard to the treatment: The counterpart of the subject whose behavior is analyzed is kept constant, i.e., the Dictators always interact with one Receiver who is not part of a group. Instead, group membership is introduced as the treatment variable.

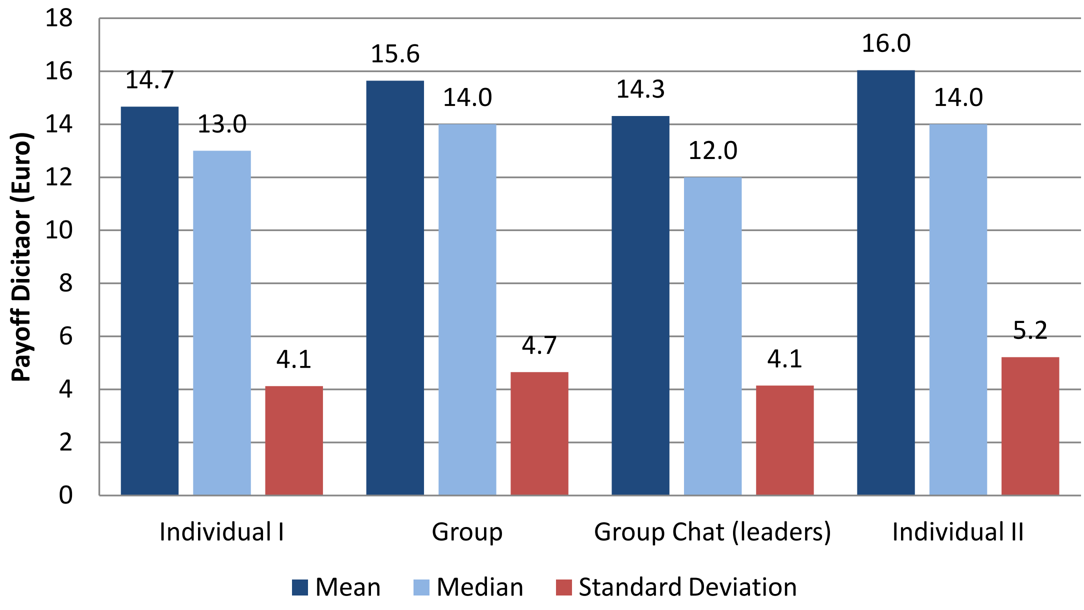
[9]The games they use are 5 two-person dictator games and 16 two-person response games with varying costs of transfer.

1.4 Results

1.4.1 Group membership

In stage *Individual I*, subjects decide individually without any induced group membership. The average amount taken by the Dictators is EUR 2.67 (median EUR 1). The decisions range from taking EUR 12 (thus taking away all of the Receiver's endowment), to giving EUR 2 to the Receiver with an overall standard deviation of the Dictators' payoff of 4.12. Clearly, the behavior formulated in hypothesis *selfish* can be rejected[10]. In all four stages, Dictators share the endowments to some degree. This behavior is in line with existing Dictator Game studies and can be explained with a heterogeneous population of agents, where some individuals have some kind of other-regarding preferences while others follow the classical, perfectly selfish payoff maximization[11].

Figure 1.2: Aggregate Dictator decisions



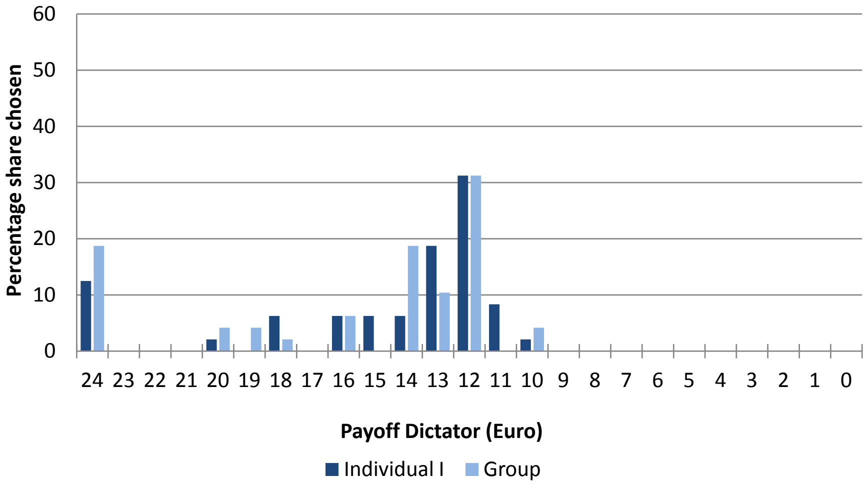
In stage *Group*, Dictators keep on average EUR 15.6 (median EUR 14) for themselves (see Figure 1.2). They, thus, take away roughly one Euro more from the Receivers when

[10]In all stages, the difference between taking everything and the observed behavior is significant at the 1% level, using both a Wilcoxon test for a single sample or a simple t-test.

[11]This well established incorporation of other agent's payoffs into the utility function can be modeled using several different approaches. These include Rabin (1993), who add a notion of fairness to standard game theory, or Bolton and Ockenfels (2000) and Fehr and Schmidt (1999) who assume inequality aversion of the subjects. However, this paper does not focus on the modeling of the individual behavior but on the potential change in behavior caused by group membership.

they act for the group than when they act alone. This change in behavior implies that hypothesis *Group* can be rejected at the 10% level (p = 0.075 using a two-sided Wilcoxon Signed Rank test). Being part of a group already changes subjects' behavior towards being more selfish and less other-regarding. A closer look at the data reveals that 47.9% of the Dictators did not change their behavior when becoming part of a group, 35.4% took more away from the Receivers while the remaining 16.7% took less away. The distribution of decisions is very similar with two clear peaks at the equal split and the perfectly selfish decision (Figure 1.3).

Figure 1.3: Distribution of Dictator decisions, stage *Individual I* and *Group*

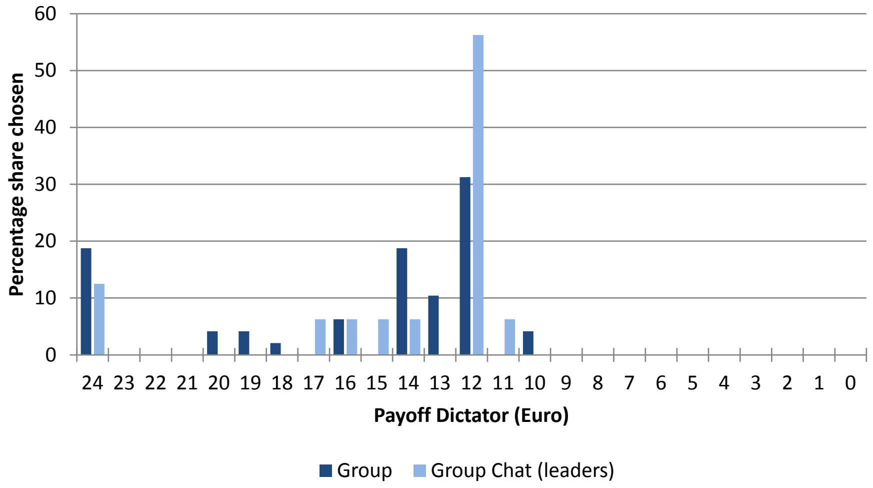


Stage *Group Chat* introduced the opportunity to chat as an additional component of group membership. The average amount the Dictators now kept for themselves was EUR 14.3, compared to EUR 14.67 in stage 1 and EUR 15.6 in stage 2. First, only the decisions by the predetermined group leaders are taken into account: The difference in decisions both between stages *Individual I* and *Group Chat* and between stages *Group* and *Group Chat* are insignificant[12]; therefore hypothesis *GroupChatA* can be rejected. Being in a group with communication did not change the leaders' behavior compared to the same subjects acting totally alone. Looking at the individual change in behavior, 44% of the subjects made the same decision in stage *Group Chat* as stage *Individual I*, while 25% took more away and 31% took less away. This balanced behavior explains the result that no difference in the aggregate behavior between these two stages is observed. The distribution

[12]Comparing stages *Individual I* and *Group Chat* results in a p-value of 0.493; comparing stages *Group* and *Group Chat* results in a p-value of 0.866. Both times a Wilcoxon Signed Rank test is used.

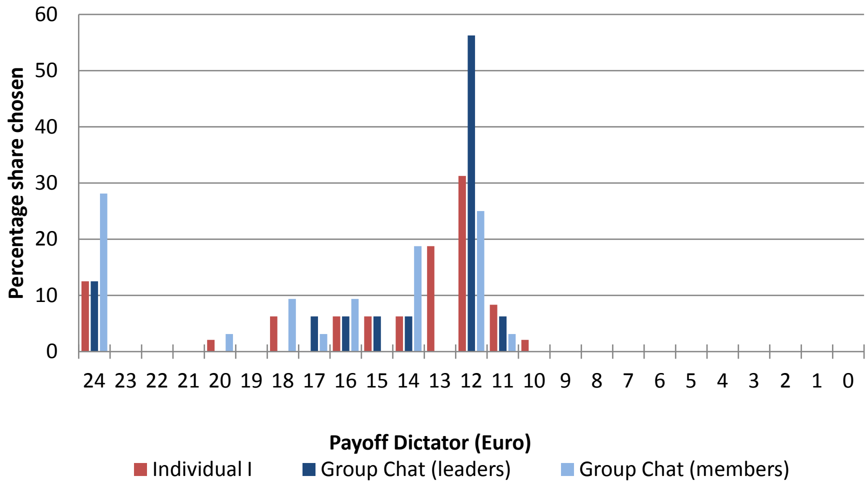
of Dictator decisions now moves away from the selfish distributions towards the equal split, with more than 55% of groups choosing this outcome (see Figure 1.4).

Figure 1.4: Distribution of Dictator decisions, stage *Group* and *Group Chat*



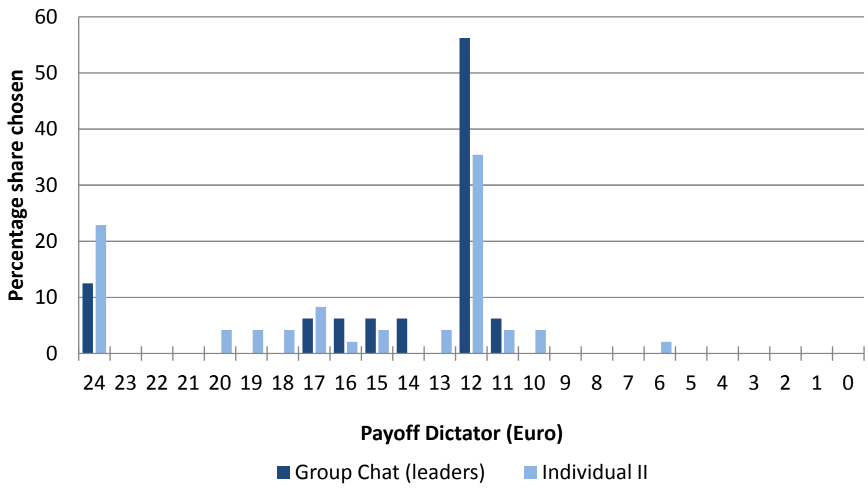
Hypothesis *GroupChatA* can be analyzed further by looking at the reaction of the leaders to the group members' preferences. As a proxy for the group members' preferences, one can look at their individual decisions in stage *Individual I*. 44% of the group leaders actually decided differently in stage *Group Chat* after the chat compared to stage *Individual I*. Of these changes, 78% changed in the direction of the group members' median decisions from stage *Individual I*. This may serve as an indicator that either the social comparison to the members of their own group or the arguments presented during the discussion serves as a motivation for some of the leaders to change their decision. However, this is unlikely as the majority of the group leaders do not change their decisions from stage *Individual I* to stage *Group Chat*. Of these, 71% encountered a group median different from their own preference. Consequently, they choose to stick to their original decision in spite of being confronted with arguments for other decisions and a social comparison which deviated from their own decision. Only two group leaders became more other-regarding despite their group members displaying preferences for more selfishness, while a more selfish group median always leads to more selfish decision of the leader. This indicates that following the group's social norm is easier for the leaders when this means a higher own payoff. This is reminiscent of the idea in social preferences where an upward deviation in payoff gives less disutility than a downward deviation (Fehr and Schmidt, 1999).

Figure 1.5: Distribution of Dictator decisions, group leaders and members



The group members hypothetically take away on average EUR 3.33 (median EUR 1) in stage *Group Chat* compared to an average of EUR 3.25 (median EUR 1) in stage *Individual I*. Thus, hypothesis *GroupChatB* can be rejected as the group members do not behave differently after the group communication than when they decided individually. The group members where not influenced by the social norm expressed in the group or by the arguments exchanged in the discussion, even though a change in their decision would have been costless to them, as the binding decision was made by the leaders.

Figure 1.6: Distribution of Dictator decisions, stage *Group Chat* and *Individual II*



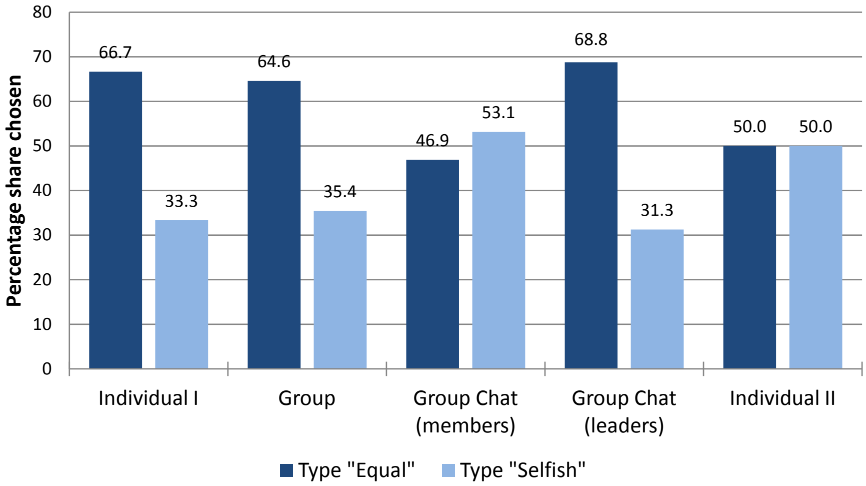
In stage *Individual II*, Dictators again acted individually and on average take away EUR 4 (median EUR 2) from their Receivers. This is significantly different (p=0.019, Wilcoxon Signed Rank test) from their behavior in stage *Individual I*. Therefore, hypothesis *SCT* can be rejected. While there is a difference, the behavior does not move into the direction of the group decision, thus hypothesis *PAT* can also be rejected. The behavior in stage *Individual II* is not significantly different from the behavior in stage *Group* (mean EUR 3.6 taken away, median EUR 2), where they decided as part of a group without communication or from the behavior in stage *Group Chat*, where communication was allowed. This points to a lasting effect of the group membership and the group discussion which took place in the preceding stages. One possible explanation is that while group membership per se drives decisions in a more rational and selfish direction, the communication aspect of group membership highlights arguments which are concerned with a socially accepted decision, thus making group members on average more other-regarding. As soon as the communication is not possible anymore, subjects then fall back to their new, more selfish behavior. This becomes also clear when looking at the distribution of divisions from stage *Group Chat* and stage *Individual II* (Figure 5). The clear peak at the equal split from stage *Group Chat* gets smaller and the decisions move back towards the more selfish divisions.

Looking at hypothesis *SCT* and *PAT* from the group members' perspective, the group members took away EUR 3.3 on average in stage *Individual I* (median EUR 1) while in stage *Individual II* they took away EUR 4.9 on average (median EUR 4.5) from their assigned Receivers. Although both decisions were made individually, this difference is weakly significant ($p=0.077$) using a two-sided Wilcoxon Signed Rank test. So while they did indeed change their behavior, they did not change it in the direction of the group decision, thus again both hypotheses *SCT* and *PAT* can be rejected.

1.4.2 Individual patterns

Finally, a way of looking at the agents' behavior is a classification of their behavior into different types. Figures 1.3 to 1.5 have indicated that there are two clear peaks in the distribution of Dictator decisions throughout the stages: At the perfectly selfish decision and at the equal split. Figure 1.7 shows a classification of agents into a "selfish type" (all agents taking away at least EUR 3 from their assigned Receiver) and an "equal type" (all agents taking away a maximum of EUR 2). There is no significant difference regarding the the type distribution (using a McNemar test) between the first two stages or between stages *Group* and *Group Chat*.

Figure 1.7: Types of agents through the four stages



However, splitting the sample between the group leaders and the group members reveals a divergence in the type classification. The subjects who actually had to make the binding decisions can be classified as 68.8% "equal type" and 31.3% "selfish type", while among those subjects who only made a hypothetical decision this division is only 46.9% "equal type" and 53.1% "selfish type". This can be interpreted in a way that the group members are aware that their decision will not influence the Receivers and thus allows them to decide only according to the payoff of themselves and their group members.

1.4.3 Chat analysis

Looking at the chat content illuminates how the decision process in the group developed. Quantitatively, the average five-minute chat included 1035 characters or 36.3 messages. The minimum of messages exchanged during a group discussion was 15, while at most 65 messages were sent during the allotted five minutes. On average, 35% of the chat was contributed by the leader, 38% by the member designated "Number 1" and 26 % by the group member designated "Number 2" when measured by the number of characters. This distribution is not significantly different from an equal distribution using a Kolmogorov-Smirnov test. As the decision rule was clear for all participants and the decision situation was well known to the participants, such extensive communication was not neccessary. Still, both leaders and members used up most of the given time. For the two group members, this suggests that they anticipated that the leader may be influenced by their preferences or arguments.

Looking at the content of the chat messages, six topics can be identified that are discussed most often: fairness, own payoff, voting / compromise, decision rule, morality and the question of earning or deserving the payoff. Table 1.1 gives an overview how many groups and subjects have discussed these topics and how many messages are sent containing one or more these topics, it furthermore includes a typical example for each topic.

Table 1.1: Frequencies of topics mentioned during chat

Topic	% groups	% subjects	# messages	Example
Fairness	50	31.3	19	I would prefer a fair division of 12
Own Payoff	62.5	31.3	23	Let's make some real money here!
Voting	81.3	35.4	25	We should vote on this
Decision rule	87.5	35.4	21	Well I decide for you
Morality	43.8	22.9	15	I think this whole thing is a little mean
Deserving	12.5	8.3	5	I mean we've all had the same task

The decision rule was communicated to every group member and judging by the frequent mentioning everybody understood what it implied. Thus, all communication was only cheap talk for the leader making the final decision. Still, in most of the groups the members try to reach a consensus and talk about a possible vote or compromise. If a consensus is actually reached during the discussion, the leader always adheres to this, even if this means deviating from her decision in the previous rounds. The fact that no group leader deviated from the intention they stated during the chat might be explained with a general aversion to lying, see for example Fischbacher and Heusi (2008) or Mazar and Ariely (2006). It is interesting to note that in all but two groups it is the leader who is actively asking for the group members' preferences. However, a clear influence of the discussed topics on the final decision can not be found[13].

[13]Using a regression with dummy variables for the different categories as explaining variables, and the leader's decision as dependent variable.

1.5 Conclusion

When determining the influence of group membership on individual behavior, it is necessary to define what exactly group membership entails. This study shows that group membership induced by payoff commonality leads to more selfish behavior if not accompanied by communication. However, when communication among the group is added, the effect disappears and subjects behave no different in the group than individually. This might be one reason for the divergent results in the literature on group membership. In general, the influence of group membership in this setup is fairly weak, giving no substantial support for either Social Comparison Theory or Persuasive Argument Theory. The analysis of the chat content reveals that in spite of the clear decision rule, groups engage in extensive communication and group members try to sway the leader towards their preferences. The other side of this behavior is the leader herself, who in 81% of the groups actively seeks the input of the other group members. The different theories of Social Preferences predict that individuals care in some way for the payoff of the other agents. Here, the leader's payoff and that of the members is always equal per the hierarchical decision rule. Therefore, the leader asking for the group members' preferences shows that he is aware that they may have different utilities associated with different payoffs. The leader not only cares for the payoff the other group members receive, but for the utility they derive from this payoff.

Chapter 2

Double or nothing!? Small groups making decisions under risk in “Quiz Taxi” [14]

2.1 Introduction

Risk attitudes are an integral part of many economic decisions and thus economic models. However, the empirical assessment of risk attitudes has proven to be very challenging because many confounding factors make isolating risk attitudes difficult. Previous economic studies have used several methods, including estimation of life-cycle macro-models, self-reported measures from micro-surveys, and laboratory experiments, to quantify risk attitudes. Each of these approaches has its shortcomings which include, e.g., the failure to account for unobserved individual heterogeneity, sample selection, or lack of incentives to reveal true preferences.

Since Gertner (1993) and Metrick (1995), economic research has been using data from TV game shows which is generated in a way that lends itself to measuring risk attitudes. This paper adds to this strand of research on decision making under risk by analyzing the behavior of contestants in the German TV game show “Quiz Taxi”.[15] The “Quiz Taxi” is

[14] See Keldenich, K., M. Klemm (2011): Double or nothing!? Small groups making decisions under risk in “Quiz Taxi”, *Ruhr Economic Papers 278*.

[15] The game show “Quiz Taxi” is the German version of the English show “Cash Cab” which first aired in 2005. It was sold to over 25 countries and is still running in several of those, e.g., USA and Australia. Bliss et al. (2011) infer coefficients of absolute risk aversion from the behavior of contestants in the US version of the show.

unique among game shows because contestants do not apply actively to be on the show, which should reduce any potential selection bias. They participate in a quiz while driving in a cab. Having reached the destination, the candidates are asked if they want to play a final "master question". If they answer correctly, their winnings will double; if they are wrong, they will lose all.

The paper makes two main contributions: firstly, the group decision about this master question is used to study risk attitudes, focusing on possible sources of heterogeneity pertaining to sex, age, and group size. Secondly, the data also allows a study of group decision making by analyzing the communication which leads to the final decision. The agents in this particular setup are groups that must reach a consensus. The individuals must therefore communicate their preferences, which in most other settings remain unobservable to researchers. Here, the "black box" of decision making can be opened to some extent. As many economic decisions are made by groups rather than individuals[16] and there might be a systematic difference in behavior between groups and individuals, this is an important aspect of the data.

In line with conventional theory, the empirical analysis reveals that contestants are risk averse. Despite fairly high chances of answering the master question correctly, only one third of all groups decides to play the master question. The tendency to take the risky master question decreases with the winnings at stake and increases with the actual performance during the cab ride. With regard to observable characteristics, all-female groups appear to be more risk averse and three-person groups more inclined to take the risk. The communication analysis reveals very strong correlations between the communication content and the decision made. This suggests that contestants make rational decisions from their subjective point of view. While about half of the groups immediately agree not to play, more than half of the initially undecided groups decide to play the risky master question. The initial response therefore appears to be dictated by risk aversion, which decreases with the length of the discussion and the arguments exchanged between group members.

The remainder of the paper is organized as follows. The next section briefly discusses the empirical background of the assessment of risk attitudes, focusing on research based

[16]Examples include household decisions about consumption and savings, "virtually all significant strategic decisions by corporations" (Cooper and Kagel, 2005), central bank decisions regarding monetary policy (Blinder, 2007), or investment decisions by mutual funds (Prather and Middleton, 2002).

on TV game shows. The "Quiz Taxi" and its contestants are described in section 3. The empirical analysis and results are presented in section 4 and the conclusions in section 5.

2.2 Empirical background

A broad range of literature about the empirical assessment of risk attitudes has developed, using several different sources of data. The various approaches are plagued by different methodological problems. Empirical macroeconomic studies try to infer risk attitudes by estimating life-cycle models of consumption. Typically, such studies assume constant relative risk aversion (CRRA) utility functions and estimate the Arrow-Pratt measure of relative risk aversion. Two early examples are Hall (1988) and Attanasio and Weber (1989) who both find fairly high degrees of risk aversion. While such estimates are useful as a description of average behavior (e.g., for the parameterization of macroeconomic models), they do not say much about individual decisions made under risk. In particular, risk aversion cannot be disentangled from other preferences, e.g., for intertemporal substitution. Furthermore, a considerable degree of heterogeneity in risk attitudes has been found depending on gender (Jianakoplos and Bernasek, 1998; Bliss et al., 2011), group size (Baker II et al., 2008), or age (Donkers et al., 2001).

For these reasons, microeconomic data is needed to study decision making under risk or uncertainty.[17] Surveys often include questions on risk attitude, either in the form of hypothetical gambles and lotteries, or as direct self-assessments. Problems with using this kind of data include the subjective nature of the answers and the lack of incentives to answer truthfully. Dohmen et al. (2011) relate such data from the German Socio-Economic Panel (SOEP) to experiments with the same SOEP respondents and find that the general risk question in the survey is a valid measure of risk attitude. Their analysis suggests that most individual CRRA parameters lie between 1 and 10. They also show that risk attitude varies with age, gender, and parental background. Barsky (1997) finds evidence for relatively high relative risk aversion (CRRA parameter of 12) with a very high variance using data from a hypothetical decision over lifetime income from the US Health and Retirement Survey. In a similar approach, Donkers et al. (2001) use the CentER Savings

[17] Since "Quiz Taxi" contestants know about the design of the decision problem and must have subjective probabilities for the possible outcomes, their decisions are made under risk, not under uncertainty. Therefore, the term "risk" is used in the following (see Knight, 1921, for the seminal contribution about the difference between risk and uncertainty).

Survey with Dutch households to estimate risk attitudes. Their semi-parametric estimates reject the assumptions of expected utility theory. Instead, decisions are better described using prospect theory (Kahneman and Tversky, 1979). It is found that decisions (and thus risk attitudes) vary with age, income, and individual wealth.

The most natural approach to studying risk attitudes are probably (laboratory) experiments where subjects must make decisions under risk with real monetary consequences in controlled environments. This method alleviates the problems of subjectivity and lack of incentives but possible problems remain, including small monetary incentives, sample selection, and experimenter effects, which would all lead to a low external validity. The seminal paper in this area by Holt and Laury (2002) compares gambles over different stakes, and real gambles with hypothetical ones. They find that a significant difference between real and hypothetical gambles emerges for larger gambles, and that the coefficient of relative risk aversion increases with the stake (from 0.3 to 0.9). Baker II et al. (2008) use an experiment to look at differences between groups and individuals and find an interaction effect: When the gamble contains either very high or very low probabilities, groups are less risk averse than individuals, while there is no difference in gambles with medium probabilities. The findings of Rockenbach et al. (2007) suggest that groups are more efficient, and better at making the trade-off between risk and higher expected payoff.

TV game shows have become a popular and valuable source of data for the study of decision making under risk (for a survey, see Andersen et al., 2008). They offer clearly defined decision situations and parameters, including the decision space, the decision maker, and the stakes. Therefore, game shows can be regarded as natural experiments (see List, 2006) because conditions between subjects are kept constant, enabling the effect of one variable to be studied with less interference than with survey data. Furthermore, the stakes are usually substantial. Due to these positive attributes, Harrison and List (2004) call situations which generate such data a "serendipity". This property was first exploited by Gertner (1993) who looks at the UK game show "Card Sharks". In this show, contestants are asked to wager a part of their winnings in a situation with clear probabilities over the outcomes. The coefficient of relative risk aversion is estimated at 4.8, but behavior is found to be incompatible with expected utility maximization. In contrast, Metrick (1995) finds that players in the show "Jeopardy!" behave in an almost risk-neutral way. Hersch and McDougall (1997) also find evidence for risk-neutrality of contestants in the lottery game show "Illinois Instant Riches" which involved very high stakes from USD -10,000 to USD

15,000. Using data from another lottery game show "Hoosier Millionaire"with high stakes of up to USD 1,000,000, Fullenkamp et al. (2003) find that contestants are risk-averse, but that the degree of risk aversion varies with the stakes.

More recently, Post et al. (2008) look at the US, Dutch and German versions of the show "Deal or No Deal".[18] Similar to the study by Gertner (1993), they find behavior to be incompatible with standard expected utility theory. Instead, the data is best organized using alternatives which are path-dependent, such as prospect theory. They can explain large parts of behavior with outcomes experienced earlier in the game, even though the stakes are high and the decision situation is clearly defined. Brooks et al. (2009) show that the risk aversion of Australian "Deal or No Deal"-contestants increases with the stake, and varies with gender and age, but not with wealth.

Matsen and Strom (2010) also discard expected utility theory (but also most other competing theories) for the explanation of contestants' behavior in the Norwegian game show "Joker" which involves large stakes of more than EUR 30,000 on average.[19] Similar to the decision that the "Quiz Taxi" contestants have to make, the "Joker" contestants can stop and keep their acquired winnings, or continue to play a risky gamble. Surprisingly, only 1 of 356 players has ever stopped. Another study which is based on a setup similar to the one at hand is Beetsma and Schotman (2001) who use data from the Dutch game show "Lingo". "Lingo" contestants have to decide up to five times if they want to continue playing with the possible outcomes of doubling their winnings or losing everything, or if they want to stop playing and keep what they have won so far. Furthermore, the contestants are always two-person groups, not individuals. Beetsma and Schotman (2001) find a high degree of risk aversion in both CARA and CRRA specifications, but do not report anything about the interaction or communication between the contestants.

Hartley et al. (2006) look at the popular game show "Who wants to be a millionaire?". This study is close to the study presented here in the sense that contestants must answer knowledge questions. Hence, their subjective assessment of their own ability has to be taken into account. Hartley et al. (2006) find that a standard CRRA framework fits the data very well. The coefficient of relative risk aversion is estimated at close to unity.

---

[18]"Deal or No Deal" has become very popular among economists. See De Roos and Sarafidis (2009) and Gee (2007) for further investigations.

[19]They find evidence for systematic expectation biases and systematic calculation errors.

Another aspect pertinent to assessing risk attitudes is the question if the decision maker is a group or an individual. In "Quiz Taxi", the decision maker is always a group, either with two or with three persons. This may have an influence on risk attitude: A study by Rockenbach et al. (2007) compares groups with individuals in a portfolio selection task and finds that groups are better at making the trade-off between risk and higher expected payoff. The "Quiz Taxi" data allows studying the communication in the group, making it possible to analyze what drives group behavior. In this respect, the study presented here goes significantly beyond the work of Bliss et al. (2011) that focuses exclusively on the estimation of coefficients of absolute risk aversion for different types of groups. They report that teams perform better and are more likely to go for the risky option. The group decision appears to be based rather on the overall money at stake than on the individual stakes of each contestant involved. However, Bliss et al. (2011) do not report anything about the subjective decision making process of the contestants.[20]

2.3 Data

2.3.1 The game show

German "Quiz Taxi" was shown from 2006 to 2008 on the TV station "kabel eins" with a total of more than 500 episodes with 2 to 5 cab rides each. For the analysis presented here, over 100 publicly available episodes from 3 DVD's and the internet portal "maxdome.de" are used.[21] Overall, the set-up of the German "Quiz Taxi" is very similar to that of the US "Cash Cab" as described by Bliss et al. (2011). Notable differences are the group sizes (1 to 5 in the US compared to 2 or 3 in the German version) and the group discussions (apparently rather limited in the US, more extensive in Germany), which explains the different focus of the study presented here and that by Bliss et al. (2011).[22]

Data from TV shows has of course some limitations (see List, 2006). Most importantly, the contestants are likely to represent only a selected sample of the whole population. This is particularly true for shows that are recorded in front of a large audience and very

[20]The authors only note that "The players are given a few moments to discuss their decision." (Bliss et al., 2011, p.7), but do not incorporate this feature of the show in the analysis.

[21]For more information about the show, see http://www.kabeleins.de/doku_reportage/quiz_taxi and http://www.kabeleins.de/serien_shows/quiz_taxi/artikel/12845 (links as of February 28, 2011).

[22]Further minor differences relate to the money earned for correct answers, and special tasks like the "red light challenges".

popular (e.g., "Who wants to be a Millionaire?"), and for shows that select participants from a specified pool of candidates (e.g. holders of a lottery ticket as in "Deal or No Deal" and "Joker"). The selection of contestants for "Quiz Taxi" is less critical in this respect: firstly, contestants do not apply actively to be on the show, but rather are quite randomly selected on the street. In an interview, the host Thomas Hackenberg says that he waits at taxi stands, gets leads from local taxi dispatchers, and often just picks up people hailing a cab (the "Quiz Taxi" is disguised as a regular cab).[23] Some contestants are also recruited on the street under false pretenses, e.g., they are asked if they would participate in a social or a market research project involving a cab ride.[24]

Regardless of the way the potential contestants are recruited, almost all of them enter the "Quiz Taxi" unaware that they can participate in a game show. The real nature of the show is not revealed to them until they get into the cab. They can then decide whether they want to play or not which obviously produces some problems of selection. In the interview mentioned above, the host notes that the reasons for backing out of the cab are very diverse.[25] It is important to bear in mind that certain population groups, e.g., fairly confident, risk loving people, and more affluent people who are able to afford taking a cab, might be overrepresented by the contestants. However, judging from the appearance and behavior of the contestants, they do not appear to be a highly selected group. In any case, selection should be much less of a problem for "Quiz Taxi" than for any other game show studied before. If there is any bias at all, it should be in the direction of choosing the risky alternative. The analysis presented below can therefore be seen as conservative estimates of risk averse behavior.[26]

[23]See http://www.kabeleins.de/doku_reportage/quiz_taxi/artikel/20708 (link as of February 28, 2010). Due to the show's increasing popularity, the host even disguises in later episodes to keep the nature of the cab hidden.

[24]This information stems from two participants. One team was offered EUR 20 each for driving in a cab operated by an unemployed person, and judging the driver's performance. After the ride, they were taken back to the starting point. The other team was neither offered any money nor taken back to the start.

[25]The host speculates that some people are afraid that they might make a fool of themselves, simply do not want to be on TV, or are just accompanied by the wrong persons. In two episodes, people who declined to play were shown. Their reasons for not participating were time constraints. Unfortunately, we do not have any concrete information about the share of participants who decline to take part.

[26]There is one additional layer to the selection problem. The producers might only air episodes that they think are interesting enough. This seems to be the case for the cab rides that were chosen for the DVD's, but not for those from "maxdome.de". The analysis therefore includes a dummy identifying 14 of 146 observations that come from the DVD's. If such a selection process were important, estimates would again be biased into the direction of risk-loving behavior. From a producer's point of view, this would also be highly cost inefficient.

Secondly, the framing of this field experiment should also be much less critical than in other quiz shows. Contestants are not placed in a TV studio but in the more familiar surroundings of a cab. In addition, there is no studio audience but only the host of the show, and the show is neither aired on one of the five biggest German TV stations nor at prime time. Bliss et al. (2011) also note that contestants cannot prepare for this particular quiz in advance.

Figure 2.1: Screenshot of "Quiz Taxi"

Source: kabel eins, Mondo Entertainment: Quiz Taxi DVD 1-1 (2008)

Figure 2.1 provides a screen shot of two "Quiz Taxi" contestants. The contestants are always groups of two or three people playing together.[27] The quiz consists of general knowledge questions (culture, history, sports, etc.) which are posed while the contestants are driven to their destination. Answering the questions correctly yields EUR 50 for the first three questions, then EUR 100 and after the 10th question EUR 150.[28] Missing a question leads to the loss of one "life", losing three lives means that the contestants lose all the money they have earned so far and have to leave the cab even if they have not reached the destination yet. There are also two "wild cards": The contestants can call somebody or they can ask passersby for help.[29] If the contestants manage to reach their

[27]This constitutes a major difference compared to the US version where group sizes of 1 to 5 persons enable Bliss et al. (2011) to investigate differences in risk attitudes between individuals and small groups.

[28]At the beginning of the first season, the first 5 questions were worth EUR 50, and each of the following EUR 100. Since in these episodes no master question was offered, they are disregarded for the analysis presented here.

[29]In later episodes, contestants can win back the previously used passersby-wild card by successfully solving an additional question or task when the cab has to stop at a red traffic light.

destination without losing all three lives, the host shows them the money they have earned so far and asks them if they would like to play the “master question”. If they answer this final question correctly, their winnings are doubled, otherwise they lose everything. For economic research, this is the most interesting part of the show because it enables the decision to choose this risky option instead of leaving the cab with the acquired money to be studied. To this end, observable characteristics of the cab ride, the contestants, and the questions are recorded: city, distance from start to destination, group size, a contestant’s sex, age[30] and possibly migration background[31], number of questions, final stake, share of correct and known answers (no wild card used, no obvious guesses), lives lost, use of wild cards.[32]

The main innovation of this paper is that it does not stop at these observable characteristics. The group discussion is transcribed as soon as the host asks the contestants if they want to play the master question until the final decision is reached. This allows aspects of the communication to be incorporated into the analysis, including the discussion time and content, initial opinions and discussion shares of group members. Since the process of recording this data is subject to measurement error, all information obtained from the transcripts was cross-checked. Furthermore, controlling for the person who coded the data does not have any influence on the obtained estimation results presented in section 2.4.[33]

2.3.2 Descriptive statistics

In total, 367 “Quiz Taxi” rides are publicly available. After excluding all rides from the beginning of the first season which were played by different rules without a master question, a few duplicates and a few rides with only one passenger, 256 cab rides with 562 individuals are left.[34] The analysis takes place on the group level because the final decision must be made by the group, and because the focus lies on the communication leading to that decision. The individual characteristics are hence aggregated to the group dimension.

[30]The age is estimated by appearance of the contestants if no direct information is revealed during the show. Three broad categories are used: below 30 years of age, 30 to 50 years of age, and above 50 years of age.

[31]Based on language skills and appearance.

[32]Table A.2.1 in the appendix provides an overview of the variable definitions.

[33]This is done by adding an indicator variable for the person who recorded the data to the regressions. The results never change. The dummy variable itself is always insignificant.

[34]The data also includes one ride which was available on the “kabel eins” website for some time. There were 5 duplicates, and 5 rides with only 1 passenger (including one celebrity).

Of the 256 groups, 43% do not reach the master question, 37.5% reach the master question and choose not to answer it and the remaining 19.5% choose to answer the master question. Table 2.1 gives the descriptive statistics by the success of the groups. Those who do not reach the final destination have of course answered fewer questions correctly and thus earned less money (which they cannot keep anyway). With regard to group characteristics, all-female groups reach the master question less often. None of the other group characteristics (size, age composition, migration background) seem to make a difference.

Looking only at the groups who reach the master question, the average stake for those who play the master question is EUR 771 (with a standard deviation of EUR 222) while those who decline to play have earned on average EUR 833 (standard deviation of EUR 255). Thus, a sizable range of stakes exists to test for a possibly stake-dependent decision. The two groups do not appear to differ much with regard to all other characteristics, including the share of questions answered correctly or known.

Table 2.2 provides more detailed information, splitting the groups that reach the master question according to their initial opinion about playing it. These descriptives already point toward a certain degree of risk averse behavior by the contestants. While half of the groups are from the very beginning of the discussion against playing the master question (only one group then played nevertheless), only 5.5% directly agree to continue playing (all did so). Among the initially undecided groups, 63% finally decide to play the master question. In total, 49 groups (33.6% of those who reach the master question) take the risky option of either doubling their winnings or losing them completely.

With regard to group characteristics, the proportion of all-female groups is relatively low among the undecided groups. As far as age composition and migration background are concerned, no large differences emerge between the “Don’t play” and “Undecided” groups. The composition of the “Play” group is somewhat different, but given the very low number of only 8 observations, these should not be over-interpreted. Not surprisingly, there are more undecided three-person groups than undecided two-person groups. It seems like the larger groups and those that are more diverse in terms of sex and age composition are less likely to have a uniform opinion.

Apart from the discussion time, the initial tendencies of the contestants and their respective share of the discussion (measured by the number of characters attributed to each

Table 2.1: Variable means (standard deviations) by success of contestants

	All contestants	Master question not reached	Master question reached ... but not played	and played	and solved
Stake (/100)	6.85	5.17	8.33	7.71	7.87
	(2.74)	(2.14)	(2.55)	(2.22)	(2.34)
Stake per contestant (/100)	3.18	2.40	3.87	3.58	3.66
	(1.30)	(1.05)	(1.16)	(1.16)	(1.23)
Distance	4.90	5.14	4.80	4.57	4.67
	(1.43)	(1.54)	(1.37)	(1.20)	(1.24)
Number of questions	10.18	9.44	10.98	10.26	10.37
	(2.13)	(2.06)	(2.06)	(1.87)	(1.96)
Lives left	0.77	0.00	1.25	1.52	1.54
	(0.81)	(0.00)	(0.54)	(0.71)	(0.74)
Share of correct answers (in %)	76.72	66.55	83.87	85.35	85.61
	(11.28)	(7.91)	(5.54)	(7.38)	(7.83)
Share of known answers (in %)	59.48	51.65	64.71	66.67	66.70
	(14.23)	(14.04)	(10.60)	(12.41)	(12.31)
Streak of correct answers	1.55	0.00	2.55	3.04	3.20
	(2.44)	(0.00)	(2.38)	(3.21)	(3.41)
Wild card left	0.21	0.19	0.18	0.30	0.29
	(0.41)	(0.39)	(0.38)	(0.46)	(0.46)
2 passengers	0.81	0.80	0.82	0.80	0.80
	(0.39)	(0.40)	(0.38)	(0.40)	(0.40)
3 passengers	0.19	0.20	0.18	0.20	0.20
	(0.39)	(0.40)	(0.38)	(0.40)	(0.40)
Females only	0.28	0.33	0.28	0.16	0.17
	(0.45)	(0.47)	(0.45)	(0.37)	(0.38)
Males only	0.29	0.26	0.29	0.36	0.39
	(0.46)	(0.44)	(0.46)	(0.48)	(0.49)
Females and males	0.43	0.41	0.43	0.48	0.44
	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)
Young contestants (under 30)	0.38	0.39	0.39	0.36	0.32
	(0.49)	(0.49)	(0.49)	(0.48)	(0.47)
Middle age contestants (30 - 50)	0.41	0.41	0.41	0.44	0.49
	(0.49)	(0.49)	(0.49)	(0.50)	(0.51)
Old contestants (above 50)	0.04	0.03	0.05	0.04	0.05
	(0.19)	(0.16)	(0.22)	(0.20)	(0.22)
Contestants of different age	0.16	0.17	0.16	0.16	0.15
	(0.37)	(0.38)	(0.36)	(0.37)	(0.36)
Migration background	0.11	0.11	0.10	0.12	0.12
	(0.31)	(0.31)	(0.31)	(0.33)	(0.33)
DVD episode	0.09	0.08	0.05	0.18	0.22
	(0.29)	(0.28)	(0.22)	(0.39)	(0.42)
Observations	256	110	96	50	41

Source: Own calculations

contestant and the host from the discussion transcripts)[35] , the communication content was quantified by recording the main topics that were mentioned during the discussion. The six main topics were "modesty/humility", "nothing to lose", "a lot of money", encouragement or provocation by the host, and the estimated difficulty of the master question ("easy" or "difficult"). As soon as contestants talk about their modesty or humility in positive terms, this topic is coded as mentioned, i.e., the corresponding binary variable takes the value one. "Nothing to lose" is coded as mentioned if contestants say that even when answering incorrectly, they do not lose anything, which could be referred to as irrationality or a wrong reference point. "A lot of money" alludes to the contestants saying that the stake is already a significant sum of money for them. This is also coded if the contestants talk about important things they could buy with the stake. The fourth topic is only coded as mentioned if the host encourages the contestants to play the master question or provokes them. It is not coded when contestants themselves behave in this way. Finally, two variables indicate whether contestants talk about the anticipated difficulty of the master question. These are also coded if the contestants talk about their own high or low degree of knowledge. Table A.2.1 in the appendix shows typical statements for each of the six topics.

Looking at the content of the discussion presented, all six topics are discussed by a significant share of groups, particularly among the undecided groups (see Table 2.2 and Table A.2.2 in the appendix). The majority shares within the undecided groups indicate a tendency toward not playing the master question. Only 22% of these groups have a majority for playing, while 35% have an initial majority against playing. However, 63% of these groups actually continue to play (see Table 2.2).

As expected, the average discussion length is higher among the undecided groups (33.6 seconds) compared to the unanimous groups (22.2 seconds). Table 2.3 reports the results of descriptive OLS regressions of the (log of) discussion length on different sets of explanatory variables for all groups that have reached the master question (columns 1 - 6), and only the undecided ones (columns 7 - 12). For the whole sample, discussion length has a strong positive correlation with being undecided, and a slightly negative correlation with the winnings at stake in the most simple specification of column 1. These significant correlations disappear when indicators for initial majorities are included in the analysis.[36]

[35] We also used the number of words instead of the number of characters to assess discussion shares. Both measures deliver the same results.

[36] This follows naturally since these indicator variables are highly collinear with the dummy variable for being undecided.

Table 2.2: Variable means (standard deviations) by initial group opinion about master question

	All contestants	Initial opinion about master question Don't play	Undecided	Play
Played master question	0.34	0.01	0.63	1.00
	(0.48)	(0.12)	(0.49)	(0.00)
Stake (/100)	8.12	8.38	7.92	7.38
	(2.46)	(2.57)	(2.24)	(3.06)
Stake per contestant (/100)	3.77	3.94	3.59	3.69
	(1.16)	(1.15)	(1.12)	(1.53)
Distance	4.72	4.78	4.64	4.82
	(1.32)	(1.42)	(1.21)	(1.25)
Number of questions	10.73	10.97	10.55	10.00
	(2.02)	(2.09)	(1.85)	(2.56)
Lives left	1.34	1.29	1.40	1.38
	(0.62)	(0.56)	(0.68)	(0.52)
Share of correct answers (in %)	84.38	84.26	84.67	83.07
	(6.24)	(5.58)	(7.02)	(5.76)
Share of known answers (in %)	65.38	64.78	66.31	63.36
	(11.25)	(10.34)	(11.85)	(14.75)
Streak of correct answers	2.72	2.74	2.78	2.00
	(2.70)	(2.36)	(3.16)	(1.07)
Wild card left	0.22	0.14	0.29	0.38
	(0.42)	(0.35)	(0.46)	(0.52)
2 passengers	0.82	0.86	0.74	1.00
	(0.39)	(0.35)	(0.44)	(0.00)
3 passengers	0.18	0.14	0.26	0.00
	(0.39)	(0.35)	(0.44)	(0.00)
Females only	0.24	0.29	0.15	0.50
	(0.43)	(0.46)	(0.36)	(0.53)
Males only	0.32	0.29	0.34	0.38
	(0.47)	(0.46)	(0.48)	(0.52)
Females and males	0.45	0.42	0.51	0.13
	(0.50)	(0.50)	(0.50)	(0.35)
Young contestants (under 30)	0.38	0.38	0.35	0.50
	(0.49)	(0.49)	(0.48)	(0.53)
Middle age contestants (30 - 50)	0.42	0.42	0.43	0.25
	(0.49)	(0.50)	(0.50)	(0.46)
Old contestants (above 50)	0.05	0.05	0.03	0.13
	(0.21)	(0.23)	(0.17)	(0.35)
Contestants of different age	0.16	0.14	0.18	0.13
	(0.37)	(0.35)	(0.39)	(0.35)
Migration background	0.11	0.11	0.12	0.00
	(0.31)	(0.31)	(0.33)	(0.00)
DVD episode	0.10	0.03	0.17	0.13
	(0.30)	(0.16)	(0.38)	(0.35)
Discussion time (in seconds)	27.25	22.14	33.58	22.50
	(21.06)	(15.00)	(25.55)	(15.30)
Humility/modesty	0.14	0.21	0.09	0.00
	(0.35)	(0.41)	(0.29)	(0.00)
A lot of money	0.16	0.19	0.15	0.00
	(0.37)	(0.40)	(0.36)	(0.00)
Nothing to lose	0.13	0.01	0.22	0.50
	(0.34)	(0.12)	(0.41)	(0.53)
Provocation/encouragement (by host)	0.33	0.32	0.38	0.00
	(0.47)	(0.47)	(0.49)	(0.00)
Easy to answer master question	0.08	0.03	0.12	0.25
	(0.28)	(0.16)	(0.33)	(0.46)
Difficult to answer master question	0.18	0.22	0.12	0.38
	(0.39)	(0.42)	(0.33)	(0.52)
Majority pro master	0.15	0.00	0.22	1.00
	(0.36)	(0.00)	(0.41)	(0.00)
Majority contra master	0.66	1.00	0.35	0.00
	(0.48)	(0.00)	(0.48)	(0.00)
Number of topics discussed	0.71	0.66	0.71	1.13
	(0.80)	(0.69)	(0.88)	(0.99)
Number of topics pro	0.21	0.04	0.34	0.75
	(0.46)	(0.20)	(0.54)	(0.71)
Number of topics contra	0.49	0.62	0.37	0.38
	(0.67)	(0.68)	(0.65)	(0.52)
Observations	146	73	65	8

Note: Majority shares do not sum up to 100% due to indifferent contestants.
Source: Own calculations

Having a majority greatly shortens discussion time. For the sample of undecided groups, discussion time is shorter for the three-person groups which can be attributed to the possibility of a majority vote. This view is supported by the result that the significant correlation between discussion length and three-person groups vanishes with the inclusion of the majority indicators.

Looking at the discussion content, it can be seen that discussion length and the number of topics discussed are strongly positively correlated (even when controlling for initial majorities). One cannot infer from these results whether the discussion topics prolong the discussion, or whether a longer discussion gives rise to more topics. Likely, both mechanisms are present at the same time. Concerning the discussion topics, the activity of the host is significantly positively related to discussion length which suggests that he plays an active role and tries to influence the contestants. The mention of the valuation of the money at stake, and of the chances of a correct answer to the master question are also positively related to discussion length. Since these aspects are essential for the decision to be made, longer discussion can tentatively be seen as an indication of more "rational" decision making (see discussion below).

2.4 Empirical results

2.4.1 Econometric specification

"Quiz Taxi" allows one to study the decision of engaging in a risky activity, namely of choosing to play the master question instead of leaving the cab with the acquired winnings. From an expected utility point of view, contestants should determine their subjective probability P of correctly answering the master question and then continue to play if the expected utility from taking the gamble is higher than the utility they get from the winnings they have acquired so far:

$$P \cdot U(X_{play}, 2 \cdot stake) + (1 - P) \cdot U(X_{play}, 0) > U(X_{stop}, stake). \tag{2.1}$$

X_{play} and X_{stop} comprise factors other than the stake that are important for but unchanged by the decision, e.g., initial wealth, and factors that are changed by the decision, e.g., cu-

Table 2.3: Estimation results for discussion length
(Coefficients from OLS regressions; all and undecided groups)

Dependent variable:	**All groups**						**Undecided groups only**					
Log of discussion time	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Undecided to play master question	0.269**	0.071	0.314***	0.115	0.299***	0.103	–	–	–	–	–	–
	(0.119)	(0.136)	(0.112)	(0.124)	(0.088)	(0.114)						
Stake (/100)	-0.053**	-0.048**	-0.033	-0.029	-0.007	-0.003	-0.026	-0.009	0.004	0.025	0.008	0.031
	(0.024)	(0.023)	(0.023)	(0.023)	(0.019)	(0.018)	(0.043)	(0.039)	(0.038)	(0.033)	(0.035)	(0.029)
Share of correct answers (in %)	0.016	0.012	0.009	0.005	-0.003	-0.008	0.004	-0.003	0.002	-0.008	-0.011	-0.020*
	(0.011)	(0.011)	(0.010)	(0.010)	(0.008)	(0.008)	(0.016)	(0.015)	(0.013)	(0.012)	(0.011)	(0.011)
Wild card left	0.007	0.063	0.061	0.109	0.214*	0.246**	0.160	0.242	0.232	0.352**	0.344**	0.438***
	(0.139)	(0.130)	(0.130)	(0.125)	(0.117)	(0.117)	(0.192)	(0.179)	(0.173)	(0.163)	(0.154)	(0.140)
3 passengers	-0.095	-0.010	-0.132	-0.048	-0.169	-0.107	-0.363**	-0.239	-0.365**	-0.239	-0.389**	-0.299*
	(0.143)	(0.145)	(0.126)	(0.126)	(0.105)	(0.104)	(0.169)	(0.173)	(0.158)	(0.154)	(0.154)	(0.153)
Females only	-0.132	-0.086	-0.104	-0.059	-0.045	-0.022	-0.054	0.016	-0.002	0.109	0.003	0.078
	(0.139)	(0.136)	(0.127)	(0.123)	(0.110)	(0.106)	(0.239)	(0.225)	(0.211)	(0.188)	(0.175)	(0.146)
Males only	0.060	0.136	0.045	0.121	-0.034	0.013	-0.080	0.068	-0.059	0.113	-0.136	-0.013
	(0.121)	(0.116)	(0.111)	(0.107)	(0.094)	(0.089)	(0.174)	(0.164)	(0.158)	(0.150)	(0.145)	(0.142)
DVD episode	0.480**	0.314*	0.251	0.112	0.057	-0.053	0.702***	0.476**	0.279	0.029	0.165	-0.049
	(0.195)	(0.175)	(0.178)	(0.156)	(0.144)	(0.138)	(0.202)	(0.196)	(0.203)	(0.177)	(0.164)	(0.157)
Majority pro master	–	-0.649***	–	-0.593***	–	-0.324**	–	-0.617***	–	-0.519**	–	-0.252
		(0.192)		(0.172)		(0.160)		(0.212)		(0.208)		(0.193)
Majority contra master	–	-0.519***	–	-0.546***	–	-0.503***	–	-0.516***	–	-0.673***	–	-0.605***
		(0.175)		(0.187)		(0.184)		(0.183)		(0.216)		(0.220)
Number of topics pro	–	–	0.316***	0.232**	–	–	–	–	0.390***	0.239	–	–
			(0.096)	(0.112)					(0.126)	(0.145)		
Number of topics contra	–	–	0.326***	0.344***	–	–	–	–	0.315**	0.442***	–	–
			(0.087)	(0.083)					(0.138)	(0.135)		
Humility/modesty	–	–	–	–	0.055	0.123	–	–	–	–	-0.316	-0.020
					(0.149)	(0.147)					(0.200)	(0.216)
A lot of money	–	–	–	–	0.302**	0.380***	–	–	–	–	0.322	0.506**
					(0.126)	(0.125)					(0.223)	(0.199)
Nothing to lose	–	–	–	–	0.240**	0.081	–	–	–	–	0.196	0.018
					(0.112)	(0.130)					(0.161)	(0.169)
Provocation/encouragement (by host)	–	–	–	–	0.683***	0.666***	–	–	–	–	0.652***	0.569***
					(0.084)	(0.087)					(0.127)	(0.118)
Easy to answer master question	–	–	–	–	0.257*	0.237	–	–	–	–	0.332*	0.362*
					(0.151)	(0.155)					(0.188)	(0.201)
Difficult to answer master question	–	–	–	–	0.552***	0.531***	–	–	–	–	0.612**	0.557*
					(0.124)	(0.129)					(0.242)	(0.286)
Constant	2.053**	2.753***	2.203***	2.990***	2.831***	3.663***	3.117**	3.790***	2.815***	3.745***	3.750***	4.542***
	(0.879)	(0.886)	(0.786)	(0.847)	(0.644)	(0.752)	(1.219)	(1.188)	(0.992)	(1.039)	(0.928)	(0.965)
Adjusted R2	0.12	0.18	0.24	0.30	0.45	0.48	0.12	0.24	0.25	0.39	0.44	0.52
Model p-value	0.001	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.000
Observations	146	146	146	146	146	146	65	65	65	65	65	65

Notes: Table reports coefficients of OLS estimations of the log of discussion time (in seconds) on specified variables, robust standard errors in parentheses. Significance levels: *10% **5% ***1%
Source: Own calculations

riosity or reputation.[37] If contestants were risk neutral (and X_{play}, X_{stop} not too different), they would be indifferent between the two options if the probability of correctly answering the question was 50%. However, the data indicates that this probability is actually much higher: The average share of right answers is 84%, the average share of known answers is 65%, and 82% of those who play the master question answer correctly. Since only 34% of all groups play the master question (see Table 2.2), one can conclude that either contestants are risk averse or that they strongly underestimate their probability of being successful. The contestants of the US "Cash Cab" appear to be more risk loving: 45% of all groups go for the final "video bonus question" which is then answered correctly by 75% of the groups. The share of groups that make it to the final question is slightly higher in the US with 64% compared to 57%, too.

In order to empirically analyze the behavior of the "Quiz Taxi" contestants, the decision to play the master question is related to different sets of explanatory variables that refer to decision, contestant, and discussion characteristics. A similar approach is taken by Beetsma and Schotman (2001) and called an "explanatory data analysis" with "rules-of-thumb variables" (Beetsma and Schotman, 2001, section 3). The general regression framework can be written as:

$$P(y_i = 1|X_i) = G(X_i, \beta), \tag{2.2}$$

where y_i is an indicator variable that takes on the value 1 if group i decides to play the master question, and the value 0 otherwise. The vector X_i comprises different sets of the variables presented in Table 2.2. These variables include, e.g., the winnings at stake, the share of correct or known answers, sex and age composition of the group, as well as discussion length and topics.

The function G relates the explanatory variables to the final decision. The most obvious choice for the estimation model is a standard probit model $G(X_i, \beta) = \Phi(X_i'\beta)$ where Φ denotes the cumulative distribution function of the standard normal. As robustness checks, Table A.2.3 in the appendix shows results from a linear probability model (OLS-estimation)

[37] For simplicity, no differentiation is made between factors that are changed by the decision and by success: e.g., there might be different reputation effects depending on not playing, playing and winning, and playing and losing.

and a logit model for a baseline specification. The results are robust to the choice of the estimation model.[38]

The main sample that is used for the following analysis consists of the 146 groups which reach their destination and are thus asked if they want to play the master question. A group of special interest consists of the 65 groups that are initially undecided whether to play the master question or not.

As discussed in Section 3, the "Quiz Taxi" contestants might represent a selected sample of rather risk-loving individuals. While it is impossible to control for this selection process, it is possible to explicitly control for the selection process that takes place during the cab ride. As argued by Bliss et al. (2011), it is unlikely that this selection is directly based on risk attitudes. But since only contestants who did not answer more than two questions incorrectly reach the master question, they might over-represent people with a higher level of knowledge who would find it easier to answer the master question. Hence, one could presume an upward bias on the coefficient of the money at stake. The data allows one to control for this possibility by estimating a Heckman selection model. The instrument used in the first stage regression for reaching the master question is the distance from the start of the cab ride to the final destination. This information can be confidently regarded as unrelated to the knowledge or any other characteristics of the contestants because the potential candidates state their destination before the nature of the show is revealed to them.[39] In contrast, the distance to the destination should significantly reduce the chances of reaching it because the longer the distance, the more questions posed, and the higher the probability of getting three questions wrong. The results of column 2 in Table A.2.3 in the appendix strongly support this view: A 1 kilometer greater distance bears a highly significant relation to an approximately 8.5%-points lower probability of reaching the master question. The results of Table A.2.3 also show that the bias arising from this selection process can be adequately controlled for by adding the share of correct answers as a confounder to proxy for the knowledge and confidence of the contestants.

[38]Since results diverge a little bit for more extensive variable specifications and particular samples between the linear (OLS) and the non-linear models (probit and logit), marginal effects from the "correct" and more intuitive probit model are reported in the subsequent Tables. No difference between results from probit or logit models arise in any specification.

[39]According to two contestants from the show, the destination was actually determined by the producers. In this case, it is definitely exogenous to the contestants.

2.4.2 Baseline analysis

The first analysis only rests on readily observable characteristics of the decision and the group: the stake, the share of right answers given by the group during the ride, a dummy variable indicating whether the group still has a wild card left (which can then be used for the master question), the group's size and sex composition, as well as an indicator variable for rides from the DVD's. With regard to risk attitude, the stake is the most important variable. If contestants were risk averse, the probability of playing the master question would decline as the stake rises.

Equation 2.1 illustrates that the group's decision to play the master question does not only depend on risk attitude, but also critically on the subjective probability of correctly answering the master question which is unobservable. In the estimations, the share of correct answers to the previous questions is used as a proxy variable. This variable also controls for the possibility that smarter contestants have acquired more money, which could bias the coefficient of the stake upward if omitted (see discussion above). This variable probably also captures other things than knowledge, e.g., confidence that has been established during the cab ride. It is chosen over other possibilities (share of known answers, or measures only based on questions worth more than EUR 100) because it is easy to remember for the contestants and the most objective measure.[40] In addition, the "Wild card left"-dummy also proxies for the group's knowledge because smarter groups are less likely to have used wild cards. A wild card should also increase the probability to play the master question because more persons are allowed to answer.

The results of the baseline estimations are reported in the first 2 columns of Table 2.4 and Table 2.5 for all groups and the undecided groups, respectively. They clearly indicate risk averse behavior by the contestants. While the raw correlation between the stake and playing the master question is negative but statistically insignificant (column (1)), the correlation becomes significantly negative with the inclusion of control variables (column (2) of Table 2.4). In the full sample, an increase in the winnings at stake by EUR 100 is associated with a decline in the probability of playing the master question by about 4%-points, which equals more than 10% of the average sample probability of playing (34%). As expected, the share of correct answers bears a significantly positive

[40] In assessing which question was known or guessed, some subjectivity enters the data. All of these measures are also strongly positively correlated with each other.

relation to the probability of playing the master question. On average, 10 questions are played during the cab ride (see Table 2.2). Hence, one more correct answer translates into a 13%-points higher probability of continuing to play. Comparing the results for all and undecided groups, it is found that the stake matters more in the full sample and the share of correct answers more for undecided groups. This suggests that the formation of initial opinions is dominated by the money at stake, but that probabilities become more important in discussions.

Most of the other included variables are not significantly related to the group's decision.[41] This is somewhat surprising for the indicator variables for having a wild card left, and being a group of three. Both can be expected to have a positive influence because they increase the probability of being able to answer the master question by enlarging the group of people who are allowed to answer. In addition, being three rather than two persons reduces the money at stake per contestant and thus personal risk (assuming that winnings are shared equally). In contrast, Bliss et al. (2011) report that three and four person groups choose to play the final question more often than individuals or two person groups. They also argue that the group decision is based on the overall amount of money at stake, and not on the stake per contestant.

The dummy variable for DVD episodes is significant in the full sample which provides evidence that these rides are indeed strongly selected, but only half in size and insignificant for undecided groups. Among the undecided, all-female groups choose significantly less often to play the master question which is in-line with the finding of Bliss et al. (2011) and Jianakoplos and Bernasek (1998) that women are more risk averse than men. The findings of Dohmen et al. (2009), Fehr-Duda et al. (2006) and Fehr-Duda et al. (2011) suggest that this result is not only due to larger risk aversion, but also due to women weighting probabilities differently than men.

Tables A.2.4 and A.2.5 in the appendix provide several robustness checks concerning the variables included in the regressions. Adding the age composition and the cultural background of the group does not affect any estimates.[42] Using the stake per contestant,

[41]The large number of insignificant coefficients explains the high p-value of the models. However, the main interest here is not on the model as a whole but on the correlations between single variables and the group's decision, i.e., on the corresponding t-test which have a much higher power than the F-test for the overall model significance.

[42]These are the variables that potentially suffer most from measurement error because age and migration background must be guessed in many cases. Given the small sample size, they are excluded from all regressions in order to keep the model as parsimonious as possible.

the log of the stake, or adding an interaction of the stake and the share of correct answers does not substantially change any findings discussed above. However, it appears that a positive effect of a three person group is rather due to less risk per contestant than to an increased probability of answering correctly. Concerning the influence of the group size on playing the master question, one also has to bear in mind the internal decision process in the group. Assuming that a consensus decision must be reached, otherwise the master question will not be played, a larger group leads to a lower probability of playing the master question. This is due to the presence of an additional "veto player" who can prevent the choice of the risky alternative.

Using the share of known answers[43] or the final streak of consecutive correct answers before the master question as proxies for the knowledge of the group does not substantially alter any of the findings, either. The respective correlations remain positive but less strong, and are only statistically significant for the undecided groups. This suggests that it is likely not only knowledge, but also confidence and enthusiasm that positively influences the decision to take the risky alternative. In that sense, the final decision might be somewhat "path-dependent" (see Post et al., 2008, for a discussion of path-dependency in the show "Deal or No Deal").

For illustrative purposes, the relative risk aversion parameter γ of the standard CRRA utility function $U(stake) = \frac{stake^{1-\gamma}}{1-\gamma}$ is estimated via maximum likelihood estimation for the probability of playing the master question.[44] This estimation is complicated by censoring issues - contestants can only bet all of their acquired winnings or nothing (see Gertner, 1993; Metrick, 1995; Bliss et al., 2011) and the unknown subjective probabilities for answering the master question correctly. Using the share of *correct* answers as a proxy for this probability, γ is estimated at 4.8.[45] Using the share of *known* answers, γ is estimated at 2.4. For a uniform probability of 50% for all contestants, γ is estimated at 1.2. All of these estimates

[43]Known answers are all correct answers that were given without the help of a wild card, or that were obvious guesses. Some subjectivity enters here because one has to decide if the answer was a guess or not. In addition, wild cards are also used in cases where the answer is actually known but contestants are hesitant and want to back up their decision.

[44]The CRRA point estimates should not be over-interpreted since they are based on a sample of only 146 independent groups. We thank our colleague Maarten van Kampen for help with the maximum likelihood procedure, and Peter Schotman and Philip Hersch for providing us with the codes that they used in their papers.

[45]Using (out-of-sample) predictions for the probability to answer the master question correctly obtained from probit regressions with the sample of groups that play the master questions gives a similar, but very imprecise estimate.

are based on the assumption of very low wealth (EUR 20 which can be seen as the value of the cab ride) and thus present conservative estimates of risk aversion.[46] For the more realistic assumption of larger wealth, the risk aversion coefficients would be even higher (e.g., 4 to 14 for EUR 1,000). In a very similar setting, Beetsma and Schotman (2001) estimate a risk aversion coefficient of 0.42 (Beetsma and Schotman, 2001, Table). Taking this value and estimating the corresponding expected probability for a correct answer gives an estimate of 24%. By and large, our estimates are in-line with the evidence presented by Beetsma and Schotman (2001). In any case, the hypothesis of risk neutrality is clearly rejected, even if people would believe that their chances of answering the master question are very low. These results also accord reasonably well with those of Dohmen et al. (2011) based on large experimentally-validated microeconomic survey data. Compared to other game show studies like Hartley et al. (2006) or Hersch and McDougall (1997), the estimates presented here are somewhat higher, which could be due to smaller selection effects at work in "Quiz Taxi", as discussed in Section 2.2.

For comparison with the results of Bliss et al. (2011) for the US "Cash Cab", coefficients of absolute risk aversion α are also estimated ($U(stake) = 1 - \exp(-\alpha \cdot stake)$). Depending on the proxy for the subjective probabilities, α hovers between 0.003 and 0.009. In line with the descriptives discussed above, the contestants of the German version of the show display a much larger degree of risk aversion. These estimates are also considerably higher than those reported by, e.g, Gertner (1993) or Metrick (1995), but fits in with the coefficients presented in Table 5 in Cohen and Einav (2007).[47]

2.4.3 Communication analysis

While it is certainly interesting to deduce risk attitudes by looking at the relationship between choosing to play the master question and the winnings at stake or the prior performance on the show, the "Quiz Taxi" offers much more information about the decision making process of the group. As discussed in section 2.3, one can observe the communication within the group, and thus gain information about the length of the discussion,

[46] Another argument for setting the reference wealth level to zero or equal to the acquired winnings comes from prospect theory (Kahneman and Tversky, 1979) where individuals are assumed to evaluate gains and losses directly and not with regard to their overall wealth (see also Beetsma and Schotman, 2001; Fullenkamp et al., 2003).

[47] Clearly, the comparability of all of these estimates is very limited because they are obtained from very different samples using various methodologies. Therefore, the external validity of estimates of risk aversion parameters is highly questionable (see Rabin, 2000).

Table 2.4: Estimation results for decision to play the master question
(Average marginal effects from probit regressions; all groups)

Dependent variable: **Play master question (1=yes, 0=no)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Stake (/100)	-0.024	-0.044***	-0.036**	-0.020	-0.025**
	(0.016)	(0.016)	(0.016)	(0.012)	(0.011)
Share of correct answers (in %)	–	0.013**	0.010	0.011**	0.013***
		(0.006)	(0.007)	(0.005)	(0.004)
Wild card left	–	0.075	0.064	0.057	-0.018
		(0.093)	(0.093)	(0.079)	(0.081)
3 passengers	–	0.071	0.091	0.114	0.148**
		(0.106)	(0.102)	(0.074)	(0.064)
Females only	–	-0.138	-0.118	-0.132	-0.142**
		(0.095)	(0.091)	(0.081)	(0.066)
Males only	–	-0.033	-0.023	-0.025	-0.015
		(0.090)	(0.087)	(0.062)	(0.060)
DVD episode	–	0.270**	0.173	0.083	0.146
		(0.129)	(0.135)	(0.110)	(0.110)
Discussion time (in seconds)	–	–	0.005***	0.006***	0.009***
			(0.002)	(0.002)	(0.002)
Number of topics pro	–	–	–	0.306***	–
				(0.061)	
Number of topics contra	–	–	–	-0.297***	–
				(0.056)	
Humility/modesty	–	–	–	–	-0.696***
					(0.109)
A lot of money	–	–	–	–	-0.476***
					(0.095)
Nothing to lose	–	–	–	–	0.465***
					(0.064)
Provocation/encouragement (by host)	–	–	–	–	-0.225***
					(0.058)
Easy to answer master question	–	–	–	–	0.058
					(0.097)
Difficult to answer master question	–	–	–	–	-0.244***
					(0.077)
Pseudo R2	0.01	0.09	0.13	0.41	0.54
Model p-value	0.129	0.027	0.008	0.000	0.000
Observations	146	146	146	146	146

Notes: Table reports average marginal effects from probit regressions, robust standard errors in parentheses. In specification (5), 4 failures and 1 success are completely determined. Significance levels: *10% **5% ***1%
Source: Own calculations

Table 2.5: Estimation results for decision to play the master question
(Average marginal effects from probit regressions; undecided groups only)

Dependent variable: **Play master question (1=yes, 0=no)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Stake (/100)	-0.005	-0.044	-0.041	-0.031*	-0.054***
	(0.028)	(0.030)	(0.029)	(0.017)	(0.016)
Share of correct answers (in %)	–	0.030**	0.028**	0.015**	0.024***
		(0.013)	(0.012)	(0.007)	(0.007)
Wild card left	–	-0.083	-0.091	-0.064	-0.152*
		(0.131)	(0.124)	(0.096)	(0.087)
3 passengers	–	0.019	0.093	0.192**	0.269***
		(0.128)	(0.123)	(0.093)	(0.102)
Females only	–	-0.299**	-0.282**	-0.344***	-0.318***
		(0.136)	(0.142)	(0.098)	(0.091)
Males only	–	-0.112	-0.089	-0.140*	-0.131
		(0.130)	(0.123)	(0.084)	(0.081)
DVD episode	–	0.132	0.005	0.137	0.231**
		(0.164)	(0.176)	(0.148)	(0.099)
Discussion time (in seconds)	–	–	0.006**	0.007***	0.012***
			(0.003)	(0.002)	(0.002)
Number of topics pro	–	–	–	0.228**	–
				(0.113)	
Number of topics contra	–	–	–	-0.373***	–
				(0.071)	
Humility/modesty	–	–	–	–	-0.515***
					(0.098)
A lot of money	–	–	–	–	-0.576***
					(0.113)
Provocation/encouragement (by host)	–	–	–	–	-0.271***
					(0.088)
Easy to answer master question	–	–	–	–	-0.087
					(0.134)
Difficult to answer master question	–	–	–	–	-0.404***
					(0.152)
Pseudo R2	0.00	0.14	0.19	0.50	0.56
Model p-value	0.870	0.209	0.112	0.000	0.000
Observations	65	65	65	65	65

Notes: Table reports average marginal effects from probit regressions, robust standard errors in parentheses. In specification (5), 1 failure is completely determined. "Nothing to lose" not included because of perfect prediction of success. Significance levels: *10% **5% ***1%
Source: Own calculations

the arguments exchanged, and initial opinions about the choice to be made. By doing so, one can to some extent open the "black box" of the decision making-process: Variables that typically remain hidden to researchers can be included in the analysis. The indicator variables for the discussion topics can be seen as proxies for some of the factors X_{stop} and X_{play} in Equation 2.1, e.g., wealth effects or personality traits. Omitting such variables could bias estimation results.

Columns (3) to (5) of Tables 2.4 and 2.5 report the results from regressions that incorporate the additional information obtained from the group discussions for all and undecided groups. The latter especially have a strong need for communication in order to come to a mutual agreement. On average, their discussions last about 11 seconds, i.e., 50%, longer (see Table 2.2). For both samples, it is not necessarily the case that the arguments are exchanged first and then the decision is made. It is possible that the decision has already been made and the arguments mentioned only serve to support this decision. This does not pose a problem because the interest is on the observation that the decision to play the master question is associated with certain traits, beliefs or motives of the contestants. It must only be assumed that the contestants do not intentionally lie about their motives which seems very unlikely. For an overview of which groups discuss which topics, see Table 2.2; Table A.2.2 in the appendix gives additional information about the discussion topics.

The inclusion of the discussion topics and the discussion length greatly improves the overall model fit. Of all 146 groups (65 undecided groups), 102 (44) discuss at least 1 of the topics. Most of the discussion topics bear a strongly significant relation to the decision made and all of them point into the direction that one would expect. In this sense, decisions made by the "Quiz Taxi" contestants can be seen as subjectively rational because they do what they believe to be right.

Groups that say that they are humble are much less likely to play, the same is true for groups that regard their acquired winnings as a lot of money. Those who believe that they cannot lose anything are very likely to play the master question. All of these correlations are much stronger for the sample of undecided groups than for the whole sample.[48]

The discussion length always has a strongly positive correlation with the decision to play the master question, which might be due to contestants realizing after some time

[48]Having nothing to lose perfectly predicts that an initially undecided group will play, and is thus not included in the regressions. These coefficients are likely inflated due to the small sample size.

that playing the master question is an attractive option or that a longer discussion may introduce another factor that also influences the decision.[49] Table 2.3 shows that the discussion length is strongly related to the activity of the host, and the topics discussed. The activity of the host is likely endogenous, i.e., he provokes or encourages contestants that do not want to play. Omitting discussion time from the regressions, no significant correlation between the activity of the host and the final decision is found. Therefore, it can be speculated that the host makes "unwilling" contestants discuss their decision longer, and that the longer discussion increases the likelihood of choosing to play the master question, which could be seen as the "rational" decision as discussed in section 2.3.2. In this case, the host would indirectly be able to talk the contestants into taking the risky alternative.

Regarding the main variables of interest from the baseline regressions, some differences emerge once the discussion characteristics are included. The stake is still negatively correlated with playing the master questions, but significance is reduced, except for the most extensive specifications in columns (5). It now also appears to be more important among the undecided groups. The share of correct answers is now significantly positively associated with playing the master question in all but one specification. This relation remains much stronger for the undecided groups.

In contrast to the previous regressions, it now also seems as if all-female groups are much less likely to play, especially when they are initially undecided. In the full sample, all-female groups have a 14%-points lower propensity to play the master questions, which even decreases to -32%-points for the undecided groups. In the most extensive specifications, three person groups are now much more likely to play than two person groups by approximately 15%-points, and by 27%-points among the undecided.

The very high R-squared of the discussion characteristics-augmented regressions might be regarded as problematic. Therefore, all regressions are also carried out including each discussion topic separately (the dummies for the difficulty of the master question are included together).[50] The results of these regressions are reported in Table 2.6. This also allows one to see which variables impact individually on the other explanatory variables, i.e., omitting which variables would lead to an estimation bias. Most importantly, none of the previously discussed findings change substantially.

[49]This significant correlation also holds if the longest discussions are excluded from the analysis, or if the log of discussion time is used (see Tables A.2.4 and A.2.5).

[50]The mutual correlations of the discussion topics are quite low.

Table 2.6: Estimation results: Each topic separately
(Average marginal effects from probit regressions; all and undecided groups)

	All groups					Undecided groups only			
Dependent variable: Play master question (1=yes, 0=no)	**Humility/ modesty**	**A lot of money**	**Nothing to lose**	**Provo- cation**	**Difficulty of master**	**Humility/ modesty**	**A lot of money**	**Provo- cation**	**Difficulty of master**
Stake (/100)	-0.035**	-0.043***	-0.023	-0.040***	-0.032**	-0.048*	-0.053**	-0.050*	-0.043
	(0.015)	(0.015)	(0.014)	(0.015)	(0.016)	(0.026)	(0.025)	(0.026)	(0.029)
Share of correct answers (in %)	0.008	0.015**	0.009*	0.012*	0.010	0.023**	0.028***	0.033***	0.027**
	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.010)	(0.010)	(0.010)	(0.011)
Wild card left	0.069	0.015	0.061	0.020	0.064	-0.104	-0.081	-0.138	-0.118
	(0.091)	(0.090)	(0.084)	(0.092)	(0.092)	(0.108)	(0.098)	(0.120)	(0.122)
3 passengers	0.101	0.133	0.096	0.124	0.085	0.121	0.233**	0.145	0.112
	(0.090)	(0.090)	(0.090)	(0.099)	(0.099)	(0.108)	(0.098)	(0.115)	(0.128)
Females only	-0.140	-0.110	-0.131*	-0.136	-0.114	-0.291**	-0.253**	-0.291**	-0.294**
	(0.087)	(0.087)	(0.075)	(0.088)	(0.090)	(0.121)	(0.126)	(0.136)	(0.136)
Males only	-0.029	-0.000	-0.066	0.006	-0.021	-0.080	-0.164*	-0.037	-0.074
	(0.081)	(0.079)	(0.078)	(0.087)	(0.086)	(0.106)	(0.096)	(0.118)	(0.120)
DVD episode	0.228*	0.246**	0.120	0.200	0.127	0.185	0.228*	0.023	0.009
	(0.133)	(0.121)	(0.124)	(0.130)	(0.136)	(0.197)	(0.127)	(0.161)	(0.180)
Discussion time (in seconds)	0.005***	0.006***	0.004**	0.008***	0.006***	0.007***	0.008***	0.010**	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Humility/modesty	-0.670***	–	–	–	–	-0.803***	–	–	–
	(0.140)					(0.194)			
A lot of money	–	-0.515***	–	–	–	–	-0.693***	–	–
		(0.115)					(0.152)		
Nothing to lose	–	–	0.572***	–	–	–	–	–	–
			(0.112)						
Provocation/encouragement (by host)	–	–	–	-0.229**	–	–	–	-0.253**	–
				(0.090)				(0.127)	
Easy to answer master question	–	–	–	–	0.157	–	–	–	-0.063
					(0.136)				(0.193)
Difficult to answer master question	–	–	–	–	-0.177*	–	–	–	-0.255
					(0.103)				(0.182)
Pseudo R2	0.23	0.22	0.29	0.16	0.15	0.33	0.36	0.24	0.22
Model p-value	0.000	0.004	0.000	0.003	0.003	0.017	0.053	0.131	0.091
Observations	146	146	146	146	146	65	65	65	65

Notes: Table reports average marginal effects from probit regressions, robust standard errors in parentheses. "Nothing to lose" perfectly predicts success for undecided groups. Significance levels: *10% **5% ***1%
Source: Own calculations

It stands out that the discussion topics that refer to the valuation of the money at stake impact most on the other coefficients. Including the "Nothing to lose"-dummy in the regressions adds most explanatory power.[51] "Nothing to lose" is negatively correlated with the stake: Groups who believe that there is nothing to lose have acquired EUR 100 less on average (see Table A.2.2). Hence, the change in the coefficient of the stake can be seen as further evidence for increasing risk aversion with regard to the winnings at stake. But it could also suggest that failing to control for reference points results in a downward bias on the estimated coefficient of the money at stake. In contrast, the "A lot of money"-dummy, which could be expected to have similar effects because it can be seen as the opposite end of the same aspect, impacts on the estimation results rather in the opposite direction, i.e., omitting this variable creates an upward bias on "Stake".

Except for the "DVD's"-dummy, none of the other coefficients change considerably with the inclusion of the discussion topics. The changes in the coefficient of the "DVD's"-dummy further support the view that these rides represent a very selected sample of particularly interesting contestants.

So far, only the determinants of the decision to play the master question have been studied. The question whether the final decision is a good one remains open, and would require knowledge of the unknown counter-factual situation. That is, would groups that decline to play answer the master question correctly? Would groups that play the master question be better off if they had stopped? The latter question can be partly addressed by looking at the performance of the 50 groups that play the master question. The small number of observation limits the possibility of a multivariate analysis, especially because only 9 groups answer incorrectly. Table 2.7 compares variable means for the groups that answer correctly or wrongly to the master question.

82% of these groups answer correctly. Therefore, playing the master question can be seen as an attractive option. Those who give the right answer, have on average acquired more money. The difference is statistically significant between the two groups at a 10% significance level and suggests that groups with more money at stake make the better decision. If risk averse, a higher stake should lead contestants to think more carefully and to continue only if the chance of answering correctly are high. A similar argument is made by Fehr-Duda et al. (2010) who attribute increasing relative risk aversion when stakes rise

[51] In the sample of undecided groups, "Nothing to lose" perfectly predicts that a group plays the master question.

Table 2.7: Variable means by answer to master question

	Correct answer	**Wrong answer**
Stake (/100)	7.87	7.00
Stake per contestant (/100)	3.66	3.20
Share of correct answers (in %)	85.61	84.14
Share of known answers (in %)	66.70	66.51
Streak of correct answers	3.20	2.33
Wild card left	0.29	0.33
3 passengers	0.20	0.22
Females only	0.17	0.11
Males only	0.39	0.22
Discussion time (in seconds)	38.12	28.00
Humility/modesty	0.02	0.00
A lot of money	0.07	0.00
Nothing to lose	0.34	0.44
Provocation/encouragement (by host)	0.37	0.22
Easy to answer master question	0.20	0.00
Difficult to answer master question	0.17	0.11
Majority pro master	0.41	0.44
Majority contra master	0.10	0.00
Number of topics discussed	0.80	0.56
Number of topics pro	0.54	0.44
Number of topics contra	0.27	0.11
Observations	41	9

Source: Own calculations

to individuals weighing probabilities more rationally. In contrast, none of the other readily observable characteristics do not differ by much between the two groups.

Looking at the discussion-related characteristics, some more differences emerge. On average, groups that answer correctly have discussed their decision more than 10 seconds longer (difference at the edge of statistical significance at 10% level), and have exchanged more arguments (not statistically significantly different). Looking at the discussion topics also supports the view that carefully thinking about the decision pays off. None of the groups that mention “Humilty/modesty” (1 group only), “A lot of money” (3 groups) or “Easy to answer” (8 groups) and only 1 of 8 groups that mention “Difficult to answer” answer incorrectly. And only 2 of 15 groups that were provoked or encouraged by the host get the master question wrong. However, 4 of 18 groups that believe that there is nothing to lose, which is not a really good argument for playing, give the wrong answer. In sum, the evidence suggests that the discussions indeed greatly help the contestants to make the right decision.

2.5 Conclusion

This paper contributes to the literature on risk attitudes of groups by using data from the German TV game show “Quiz Taxi”. After a cab ride during which contestants can earn money by answering knowledge questions, they are asked if they want to play a final master question with which they can double their winnings, or lose everything.

“Quiz Taxi” has two main features that make it very attractive for studying risk attitudes. Firstly, the group of contestants represents a much less selected sample than in other TV game shows because the contestants do not actively apply to be on the show and do not play in front of a large TV studio audience. Secondly, the “Quiz Taxi” contestants always play in groups of two or three people. This makes it possible to observe the communication between group members, and to open the “black box” of decision making to some extent.

Overall, contestants show fairly risk averse behavior. The risk parameter of a standard CRRA-utility function is estimated at rather high values of approximately 1 to 5 for reasonable subjective probabilities of being able to answer the master question. Compared to other studies using game show data, the analysis supports findings of, e.g., Beetsma and Schotman (2001) or Dohmen et al. (2011) of rather high but not unreasonable degrees of risk aversion which is likely due to smaller selection effects in the setup presented here. Still, there is reason to believe that the Quiz Taxi contestants are on average more risk-loving than the average population. Hence, the evidence presented should be regarded as conservative estimates of risk aversion.

The regression analysis suggests that an increase in the winnings at stake by EUR 100 (the average value of one question) is associated with a decline in the probability of playing the master question by approximately 4%-points on average. A higher share of correct answers during the preceding cab ride of 10%-points (1 more correct answer) is associated with an increase in the probability of playing by about 13%-points. This positive relationship can be attributed to greater knowledge and/or more confidence of the successful contestants. It is also shown that all-female groups are less likely to choose the risky alternative, and that three-person groups are slightly more risk-loving than two-person groups. Both findings are also reported by Bliss et al. (2011) who analyze risk attitudes with data from the US version of the show called “Cash Cab”. While it remains unclear whether women are really more risk averse or underestimate their chances of answering correctly, the analysis suggests that for three person groups the lower risk per contestant weighs more heavily than the higher probability of a right answer for the final decision.

The analysis of the communication characteristics and its content shows that contestants typically make consistent decisions in the sense that they have arguments to back up their decision, but not necessarily in the sense that the set of arguments discussed is exhaustive. Not playing the master question is related to arguments of humility or modesty, the difficulty of the final question, or the large amount of money at stake. Playing the master question is related to arguments of having nothing to lose (a different reference point), or the easiness of the final question. It is shown that the inclusion of these characteristics matters for the estimation of the other coefficients. In particular, failing to account for the individual valuations of the money at stake creates an omitted variables pertaining to the coefficients of the money at stake.

The analysis also reveals that it is important to differentiate between groups that are immediately determined to stop or to continue playing (90% of these groups stop), and groups that are initially undecided. The decision to play is positively related to the discussion length, especially for initially undecided groups, which might induce more rationality into the decision making process. While the initial tendency is dominated by risk aversion with regard to the money at stake, a longer discussion brings about more arguments and enables contestants to make better informed decisions.

The behavioral analysis highlights the importance of people's beliefs, expectations, and values for decision making processes in situations involving risky choices. Future research should account for such typically unobserved personal factors, especially in the context of group decision making where arguments must be exchanged in order to reach a mutual agreement. Since the "Quiz Taxi" has been aired in many countries since 2005, this show also represents an opportunity to study cultural differences in decision making behavior under uncertainty. Compared to their US counterparts (Bliss et al., 2011), the German contestants seem to be more risk averse at first sight.

Chapter 3

Teaching in the lab: Financial incentives in the education process [52]

3.1 Introduction

The transfer of knowledge is an important part of many economic and everyday situations. Examples can be easily found in the corporate sector (when new employees need to receive firm-specific information from their colleagues), in private life (when parents give advice to their children career choice, for example), and most prominently in the education sector (which is in essence mainly occupied with the transfer of knowledge). While standard economic theory maintains that people need incentives to exert effort, little is known about the interplay of incentives and effort in knowledge transmission processes. In particular the role of monetary incentives in this context is far from clear even though there can be little doubt that financial considerations play a big role at least in professional environments. The paper presented here aims to shed some light on this aspect by analyzing the role of monetary incentives in knowledge transfer, using a laboratory experiment.

Although in general both agents in a knowledge transmission process - the provider and the receiver of knowledge - can be subject to incentive problems, it is particularly important to incentivize knowledge providers as they usually do not have an inherent incentive to spend effort on the task. New employees for example may be motivated to learn with a view to increasing their future opportunities. Their advisors on the other

[52] See Helbach, C., K. Keldenich (2012): Teaching in the lab: Financial Incentives in the Education Process, *Ruhr Economic Papers 328*.

hand may shirk, for instance because they concentrate their effort on incentivized tasks to increase their monetary payoff. In such a scenario the knowledge provider should receive a reward for her efforts that should ideally be dependent on the amount of knowledge the intended recipient actually receives. This incentive-compatible approach is however rarely used. School teachers for example often receive with a fixed wage, regardless of their students' performance[53]. Similarly, companies seldom include knowledge transfer in the criteria used for variable compensation.

In many industries, monetary incentive schemes are however a standard way to motivate employees in areas apart from knowledge transfer. Adams et al. (2009) estimate that in 2002, 33% of the Fortune 1000 companies used some kind of individual incentive. The economic argument for variable pay linked to performance is to align the interests of employers and employees (Kessler and Purcell, 1992). Existing research has shown that monetary incentives indeed affect agents' effort in many settings. Several studies have found a positive effect of performance pay on worker effort and productivity. This effect is however task dependent and Camerer and Hogarth (1999) conclude in a survey of experimental results that "The data show that incentives sometimes improve performance, but often don't" (pg. 32). In another survey Prendergast (1999) looks at the provision of monetary incentives in firms and concludes that "incentives matter" (pg. 11).

There is, however, also evidence that the introduction of monetary incentives might have adverse effects: Gneezy and Rustichini (2000) show that an extrinsic source of motivation like money might crowd out intrinsic motivation to do the task well and therefore diminish effort. As a consequence low monetary incentives reduce performance in comparison to a scenario without monetary incentives. In contrast, Pokorny (2008) finds that even low piece rate payments improve performance in a real effort experiment. Despite these diverging key results both studies provide evidence for non-monotonic effects of monetary incentives. They thus exemplify that finding the optimal incentive scheme is not a trivial exercise. This is even more so if, apart from possible crowding out effects, performance is hard to measure, as it is the case with knowledge transfer or learning success. Adams et al. (2009) give several examples where this problem leads to a worse outcome compared to a situation without conditional monetary incentives and Prendergast emphasizes that "there has been insufficient focus on workers whose outputs are hard to observe" (pg. 11).

[53]This is e.g. true for most teachers Germany, who do not receive any variable wage at all.

To the best of our knowledge, this paper is the first to analyze the actual teaching process in the laboratory. There is, however, a literature on information transmission by way of advice giving (e.g. Chaudhuri et al., 2006; Schotter and Sopher, 2007). This literature has shown that participants in experiments strongly respond to advice, even in an anonymous experimental setup. These studies differ however from the paper at hand in at least two significant ways. First, advice deals with situations such as public good or ultimatum games. Thus, individual preferences and interaction play a role and there is no such thing as correct advice. Second, the focus is the advice giving environment and incentive schemes are not alternated as treatment variables.

While there is no laboratory evidence on monetary incentives in knowledge transmission processes, recent studies from the field of education economics have provided field evidence on teacher performance pay. The most compelling evidence stems from field experiments in India (Muralidharan and Sundararaman, 2011) and Israel (Lavy, 2002, 2009). The data from these experiments reveal a significantly positive effect of merit pay systems. Figlio and Kenny (2007) as well as Woessmann (2011) draw a similar conclusion from analyzing survey data from the United States and cross-country data, respectively. Despite those positive results there is also evidence that sheds some doubt on teacher performance pay. Reback (2008), Eberts et al. (2002), and Martins (2009) find some negative effects of teacher performance pay by analyzing data from the United States, and Portugal, respectively. The study by Reback in particular provides useful insights. He analyzes the effect of the American "No Child Left Behind Act" of 2001 which penalizes schools which do not meet minimum proficiency requirements. The author finds that teachers as a consequence focus on those students close to this minimum proficiency threshold which results in an improvement only for relatively low performing students. Similarly, Glewwe et al. (2010) find that teachers focus on incentivized goals and disregard non-incentivized ones.

The brief literature overview shows that a final understanding of the effects of monetary incentives requires further research. The present paper aims to provide one further step in this direction by analyzing instructors' incentives in a knowledge transmission laboratory experiment. It contributes to the literature in various ways. First, the controlled laboratory environment reduces problems of measuring effort by providing observable outcomes. Second, the experimental setup allows a direct comparison of four different incentive schemes at reasonable costs. Third, prospective teachers can be compared to other participants[54] to

[54]See Section 3.2.2 for a detailed explanation.

draw preliminary conclusions on whether they respond differently to monetary incentives in this specific task.

3.2 Experimental design

3.2.1 Structure

All subjects in the experiment are assigned to one of two roles, instructor or pupil, which they keep throughout the whole experiment. The experiment consists of three parts. All subjects are placed in sound protected cabins with closed doors during the experiment. Subjects are allowed to make notes but are asked to use the notepad provided by the experimenters. The instructions include information about all three parts of the experiment[55].

In the first part, only subjects in the instructor role (from now on "instructors") participate. After reading the instructions, they see a presentation on two topics. The instructors can freely move through the slides of this presentation, but only have 20 minutes time to do so. The two topics used in the presentation are a self-developed card game called "Pizzabäcker"[56] and the artificial language "Lojban"[57]. These two topics have been chosen because they fulfill several requirements: No subject should have prior knowledge of them, the topics should yield enough questions to allow variance, they should be easy to translate into multiple choice questions, and sufficiently complex to explain and teach. Two topics rather than only one were chosen to decrease the influence of idiosyncratic capabilities of the participants (if, e.g., one subject is very talented at understanding languages, the addition of a second unrelated topic decreases the distorting influence of this). Both topics take up the same number of slides in the presentation, which is known by the participants beforehand. In addition to the presentation itself, instructors also receive additional material to help them understand the topics better: a complete set of cards for the card game and a sheet containing the alphabet and one sentence for the language. During the presentation for the instructors, the subjects in the pupil role (from now on "pupils") also enter the laboratory. Every pupil is randomly assigned to one instructor. The pupils are

[55] A set of complete instructions can be found in the Appendix.

[56] The card game's name translates to "Pizza baker" and was adapted from the game "Dia de los muertos". The original game was created by Frank Graham, who graciously allowed us to use it.

[57] See http://www.lojban.org for a short introduction.

given the same instructions as the instructors (including the material belonging to the card game and the language). While the pupils read those instructions, the instructors finish the presentation and then have 10 minutes time to prepare for the second part of the experiment.

In the second part, the actual teaching takes place. For each instructor-pupil pair a one-way video conference is established, such that the pupil can see and hear the instructor but not vice versa. This one-sided mode of communication was chosen to rule out the possible influence of the pupil on the quality of the teaching. If a pupil is very smart or motivated, she can positively influence the teaching quality and effort by, e.g., asking good questions. As the aim of the design is to isolate the effects the treatments have on the instructors, this is not desirable. The connection between pupil and instructor lasts for 10 minutes, and the resulting video is recorded with the knowledge (and prior consent) of the subjects. Before the 10 minutes start, instructors are given some extra time to calibrate the position of the camera, their headset and the volume. During the 10 minutes, pupils can only press a button which indicates to the instructor that they can see and hear her properly. Instructors are not given any encouragement on how to spend the 10 minutes, they can e.g. easily avoid teaching altogether and concentrate on their notes instead (some instructors do in fact point the camera away from them so their pupil cannot see them).

In the third and final part, instructors and pupils all have to answer the same 30 multiple choice questions (15 for each topic), where four possible answers are given for each question and exactly one answer is correct. Each question is displayed for 40 seconds on the screen, the subjects cannot speed up the questions or return to older questions. This procedure is described in the instructions for everybody. After the questions are answered, every subject is informed how many questions and which questions he has answered correctly. In the treatments where instructors' payoff depends on the pupils' answers, the instructors are additionally informed about the number of correct answers of their assigned pupil.

After the experiment itself, subjects fill out a questionnaire with questions about demographics (sex, age, study length, study subject), their school grades (last math grade and last German grade), some personality measures, teaching experience, general card game experience, risk attitude, and if they knew "Lojban" or the card game before the experiment.

3.2.2 Treatments

Pupils always receive the same payoff: For each correctly answered question, they receive EUR 0.75. Instructors also receive EUR 0.75 for each question they answer correctly, but may get an additional payoff depending on the treatment they are in. In the treatment *Fix*, instructors receive an additional payoff of EUR 4.50 no matter how good their pupil is. In the treatment *Linear*, instructors receive an additional EUR 0.30 for each question their pupil answers correctly. In the treatment *Bonus*, they receive an additional EUR 9.00 if their pupil has at least 15 correct answers. For the treatment *Tournament*, the instructor-pupil pairs are each randomly assigned to groups of three pairs. The instructors then receive an additional EUR 13.50 if their pupil is the best one in their group. No show-up fee was paid in any of the treatments.

The amount of money given in the individual treatments was calibrated in such a way that the ex post payoff average for the instructors remains roughly constant in all treatments[58]. Otherwise, a difference in teaching between the treatments might be due to the higher amount of money earned and not the conditionality of the payoff.

In addition to the payoff scheme, a second treatment dimension is the subject pool used in the instructor role. To approximate actual teachers, students of educational science who want to become teachers are used[59]. In the following, the abbreviation "EDU" is used for those students, while "Non-EDU" designates other students.

Combining these two treatment dimensions results in 8 treatments. Table 3.1 shows the number of observations for each treatment.

Table 3.1: Number of observations by treatment

	Incentive Structure			
Subject Pool	Fix	Linear	Bonus	Tournament
EDU	24	23	22	22
Non-EDU	23	24	24	32

[58]This was done by first running the *Linear* treatment and then choosing the figures according to the results.

[59]The German system of teacher education is set up in a way that studying a subject to become a teacher is a distinct degree and university career from studying the subject per se (These students are called "Lehramtsstudenten"). For example becoming a chemistry teacher means studying "chemistry to become a teacher", and not the regular subject chemistry. This makes it possible to claim that a large share of the subjects used in the experiment will indeed become teachers.

3.2.3 Procedures

The experiment was conducted computer-based and took place at the "Essen laboratory for experimental economics" (elfe) at the University of Duisburg-Essen in May and November 2011. Participants were recruited via the program ORSEE (Greiner, 2004) and the attached subject pool[60]. To program the experiment, the software z-Tree (Fischbacher, 2007) was used. For the video conference, a customized version of the software Vivicom was used. In total, 34 sessions with up to 12 subjects each were conducted, leading to a total of 392 participants[61]. The participants were all students from the University of Duisburg-Essen. To avoid a confounding of possible treatment effects with gender effects, all treatments included a balanced ratio of all possible gender pairings (male instructor/male pupil, female instructor/male pupil, male instructor/female pupil, and female instructor/female pupil)[62]. The experiment lasted about 90 minutes for the instructors and 60 minutes for the pupils. Average payoff was EUR 20.23 (minimum EUR 9.75, maximum EUR 30.75) for the instructors and EUR 12.13 for the pupils (minimum EUR 4.50, maximum EUR 18.75). Subjects were paid out one after the other to preserve anonymity; the instructors were paid out before the pupils.

3.3 Hypotheses

Instructors in the present experiment can invest the restricted time during a session in learning and teaching. Even if both efforts are not completely separable it seems plausible to assume that instructors face a trade-off between teaching and learning due to both the restricted time available and the complexity of the content. From a theoretical point of view it is thus evident that homo oeconomicus like instructors do not exert teaching effort in the treatment *Fix*. In contrast, any of the implemented incentive schemes shifts priorities towards teaching under very weak assumptions[63]. If we define the share of knowledge an instructor is able to transfer to her pupil as "number of correct pupil answers" divided by

[60] Table A.3.1 in the appendix shows some descriptive statistics of the participants

[61] Of those, 4 subject pairs experienced technical problems and therefore are excluded from the analysis.

[62] Due to the aforementioned exclusions, the balancing was not perfect.

[63] Of course, one can think of preferences or expectations that e.g. let instructors shy away from competition. These would need to be very strong, however, to prevent every kind of priority shifting through the incentive schemes.

"number of correct instructor answers" for each pupil-instructor pair, we can thus formulate hypothesis 1 as follows:

The lowest share of knowledge transmitted occurs in the treatment Fix.

This hypothesis is tentatively supported by previous research. Even though it is difficult to compare both the task as well as the incentive schemes in this paper to the previous literature, there is some evidence that monetary incentives can increase performance (e.g. Bull et al., 1987; van Dijk et al., 2001; Lazear, 2000; Shearer, 2004)[64]. These papers furthermore seem to point at better performance under tournament than under piece rate incentives. From a theoretical point of view, however, this result depends on preferences and expectations of agents and on the specific setup of the tournament[65].

As for the two different subject pools, a priori the rationale still holds: If teaching causes disutility, a monetary incentive for it will increase effort. While it can be argued that teachers are maybe intrinsically motivated and therefore should not react to monetary incentives, the before mentioned literature on teacher performance pay indicates that they do. This leads to hypothesis 2:

There is no difference in the incentive reaction between the subject pools EDU *and* Non-EDU.

3.4 Results

3.4.1 Treatment effects

Figure 3.1 shows the average share of knowledge transmission for all eight treatments. Looking first at the Non-EDU subjects, the highest average ratio is 0.92 in the *Bonus* treatment, followed by the *Linear* (0.89), the *Tournament* (0.88), and the *Fix* (0.75) treatment. Testing these differences pairwise reveals a significant difference between the *Fix* treatment and all other treatments ($p<0.05$ for all pairwise tests[66]). The treatments

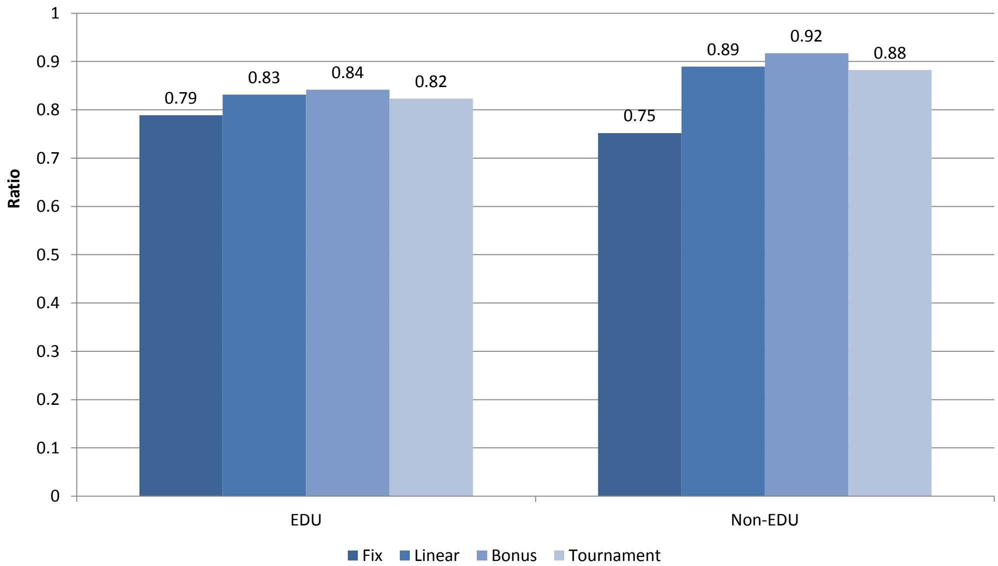
[64]Note, that a lot of experimental studies discuss pay for performance that is added on top of a fixed wage. See Camerer and Hogarth for an overview. In contrast, participants in the treatment *Fix* of the current paper receive a higher fixed payoff than participants in other treatments. Against the background of the literature on efficiency wage theory and gift-exchange games (c.f. Akerlof, 1982; Akerlof and Yellen, 1990; Fehr et al., 1998) this aspect hinders a direct comparison.

[65]See for example Harbring and Irlenbusch (2003) for an experimental investigation of the latter aspect.

[66]If not indicated differently, exact two-sided pairwise Wilcoxon rank-sum tests are used.

with a conditional monetary incentive (*Linear*, *Fix*, and *Tournament*) are not significantly different from each other in terms of the ratio (all tests yield p>0.1). Therefore, the formulated hypothesis cannot be rejected for the Non-EDU subjects.

Figure 3.1: Average ratio of correct answers pupil/teacher by treatment



Now looking at the EDU subjects, the ordering of treatments remains the same, but the difference between the *Fix* treatment and the other treatments is less pronounced: The highest ratio was achieved under the *Bonus* payment system (0.84), followed by the *Linear* (0.83), the *Tournament* (0.82), and the *Fix* (0.79) payment scheme. These differences are not significant (all tests yield p>0.1). Therefore, the hypothesis can be rejected for the subjects who study to become teachers: The payment scheme has no influence on their teaching performance[67].

Comparing the two different subject pools with each other for each treatment separately reveals that there are no significant differences between prospective teachers (the EDU subjects) and students who do not want to become teachers (the Non-EDU subjects)[68]. Still, as shown previously, the reactions of the teachers on the different payment schemes are different for both groups.

[67] Figures A.3.1 and A.3.2 in the appendix show the ratios of transferred knowledge for each of the two topics separately. Qualitatively, the results are the same as for both topics combined. The Non-EDU subjects have the lowest ratio of transferred knowledge in the *Fix* treatment. The difference between *Fix* and the other treatments is weakly significant (p<0.1) for the card game, while for the artificial language only the test of *Fix* vs. *Bonus* yields a weakly significant result. There are no systematic (or significant) differences for the EDU subjects.

[68] All tests yield p-values > 0.1.

Figure 3.2: Average number of correct teacher answers by treatment

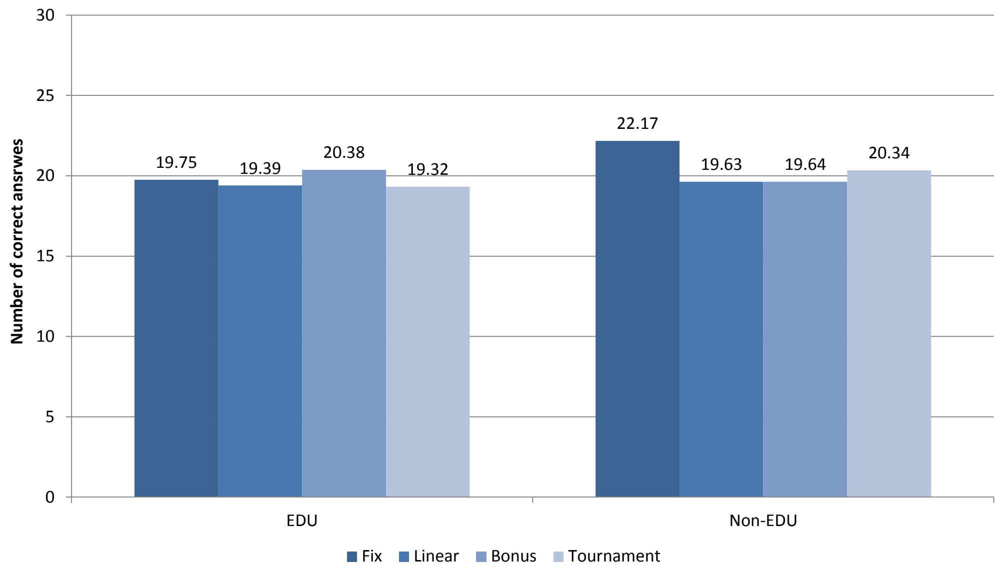
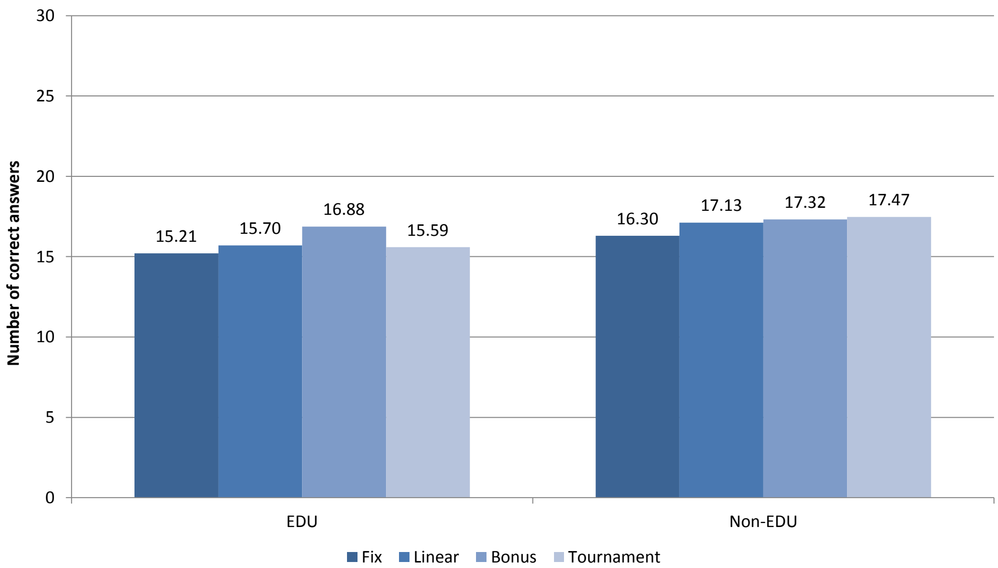


Figure 3.3: Average number of correct pupil answers by treatment



As the ratio of knowledge transfer is a combined measure of the correct answers by an instructor and her pupil, one can take a look at what drives the results in the ratio by analyzing these two underlying numbers shown in Figures 3.2 and 3.3. Again first looking at the Non-EDU subjects reveals that the difference in ratio is driven mainly by the instructors' correct answers: The instructors have the highest number of correct answers in the *Fix* treatment[69]. The pupils have the lowest number of correct answers in the *Fix* treatment, this is however not significant[70]. One possible explanation for this result is that instructors have some limited capacity of effort which they can divide between two activities: Preparing to answer their own questions and preparing to teach their pupils. If the latter is incentivized, instructors increase their effort for this and therefore must decrease their effort for the preparation for their own questions, leading to the results presented here[71]. For the EDU-students, there are no significant differences for neither the instructors' nor the pupils' correct answers.

3.4.2 Video analysis

To open up the "black box" of the teaching process, the instructor videos are looked at directly and coded along several dimensions.[72] The analysis here is focused on the objectively codable variables: The time actually used to teach is recorded, starting with the first remark with content (thus excluding "Can you hear me?" or similar). Additionally, the time(s) when the topic is switched and the first topic are recorded, such that one can determine how much time was used for each topic. Furthermore, the type of pronoun used by the instructor (i.e. if the instructor refers to herself as "I" consistently or if she includes the pupil by using "we" at least once) and the type of address (i.e. if the instructor directly addresses the pupil at least once or if he uses an indirect address) are coded. Finally, it

[69] $p<0.1$ for the comparison *Fix* vs. *Tournament*, $p<0.05$ for the other comparisons.

[70] All pairwise comparisons $p>0.1$.

[71] Note that the number of pupils' correct answers can be seen as a function of instructors' teaching effort and pupils' effort while the number of instructors' correct answers is only a function of their studying effort. Consequently, an effort shift because of changed incentives has stronger effects on the instructors' answers than on the pupils' answers if pupils' effort is assumed to remain constant.

[72] Table A.3.2 in the appendix gives an overview of all coded variables. The more subjective variables are not used in the analysis. The coding of these variables proved to be too unreliable even though all coding was done by two different coders. This is especially true for the variables which measure how good each question is answerable with only the information given by the instructor. Of the 30 questions coded this way, in about one third the two coders had discrepancies more than 40% of the time.

is recorded if the instructor points out to the pupil that she has the same material at her disposal.

Table 3.2 shows the instances of the variables "We" and "Direct Address" split along the two different subject pools. An exact χ^2-test shows that the difference between the instructors who study to become teachers and other instructors in their use of these two variables is statistically significant ($p<0.05$)[73]. A possible explanation for this behavior lies in the different perceptions prospective teachers might have concerning their role as instructors in the experiment. They see themselves as "proper" teachers and still try to keep their pupils engaged, leading to the use of a direct address. The other students however may not perceive themselves in a substantially different role than their pupils, thus using the pronoun "We" significantly more often[74].

Table 3.2: Use of direct address and we

	Direct address	We
EDU	68	60
Non-EDU	84	41
All subjects	152	101

Figure 3.4 shows the average time instructors use to teach the different topics. Instructors spent more time for the topic "Pizzabäcker" - which was also the one they saw first in the presentation - than for the topic "Lojban". However, there are no significant differences between treatments, regardless of the subject pool. The observed differences in outcome for the Non-EDU teachers are therefore not caused by simple quantity of teaching, but must have their causes in the manner of teaching.

3.4.3 Regression

A different way to look at the data is by using a regression. Table 3.3 shows the results of a simple OLS regression with the share of transferred knowledge as the dependent variable. The regression is performed for the whole sample, just with the EDU students and just with the Non-EDU students.

The results partly confirm those from the non-parametric tests above: The treatment dummies for the *Linear* and *Bonus* treatment have a significantly positive effect (with the

[73]Comparing these variables along the different incentive schemes, however, does not reveal any systematic differences.

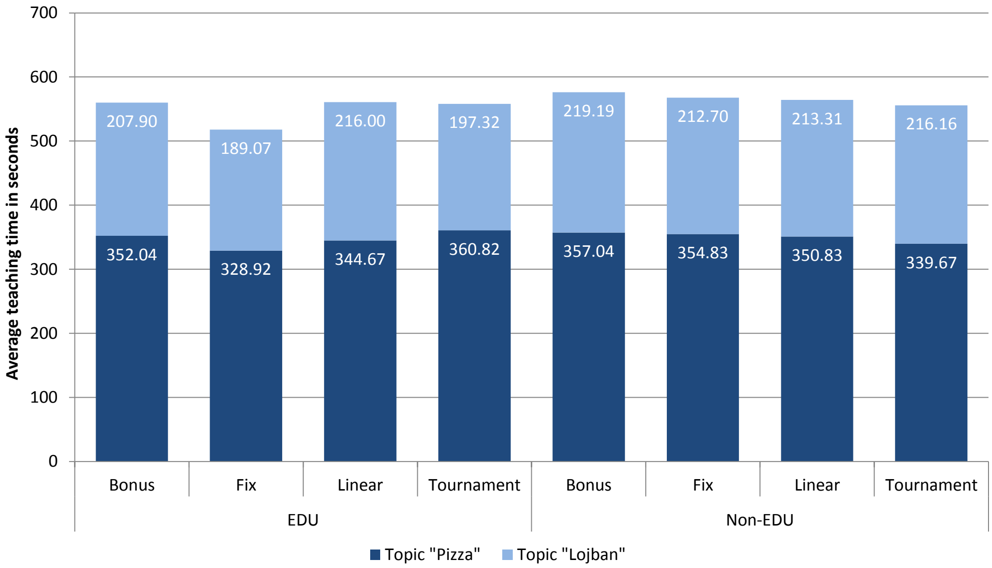
[74]Note that the type of address does not influence the results in terms of transferred knowledge.

Table 3.3: OLS estimation results for ratio of transferred knowledge

Dependent variable: **Ratio of transferred knowledge**	**All pairs**	**Only pairs with Non-EDU instructor**	**Only pairs with EDU instructor**
(Constant)	1.170***	1.562***	0.990**
	(0.000)	(0.002)	(0.034)
Linear treatment	0.164**	0.160**	0.039
	(0.019)	(0.033)	(0.575)
Bonus treatment	0.166**	0.141*	0.091
	(0.020)	(0.067)	(0.201)
Tournament treatment	0.100	0.095	0.025
	(0.128)	(0.168)	(0.724)
Linear x EDU	-0.132	–	–
	(0.168)		
Bonus x EDU	-0.090	–	–
	(0.355)		
Tournament x EDU	-0.072	–	–
	(0.457)		
Instructor EDU	0.016	–	–
	(0.823)		
Instructor female	0.011	0.008	0.019
	(0.752)	(0.866)	(0.716)
Pupil female	0.007	-0.002	0.057
	(0.837)	(0.971)	(0.315)
Teaching time	0.000	0.000	0.000
	(0.454)	(0.792)	(0.379)
Pupil material	-0.085**	-0.085	-0.071
	(0.022)	(0.130)	(0.188)
We or I	0.023	0.090	-0.057
	(0.519)	(0.103)	(0.290)
Direct Address	-0.055	-0.031	-0.065
	(0.205)	(0.674)	(0.244)
Card game experience instructor	-0.026	-0.047	-0.010
	(0.471)	(0.383)	(0.846)
Card game experience pupil	0.051	0.040	0.049
	(0.164)	(0.458)	(0.338)
Instructor's age	-0.006	-0.017**	0.006
	(0.284)	(0.040)	(0.431)
Pupil's age	-0.003	-0.009	-0.004
	(0.526)	(0.209)	(0.535)
Instructor's math grade	0.003	-0.015	0.023**
	(0.651)	(0.177)	(0.044)
Instructor's German grade	0.001	-0.004	0.010
	(0.888)	(0.794)	(0.438)
Pupil's math grade	0.000	0.000	-0.003
	(0.952)	(0.977)	(0.785)
Pupil's German grade	-0.007	-0.004	-0.031*
	(0.477)	(0.797)	(0.055)
Observations	180	92	88

Notes: Fixed treatment is the baseline for the treatment dummies. P-values in parentheses.
Significance levels: *10% **5% ***1%.
Source: Own calculations.

Figure 3.4: Average time used to teach each topic by treatment



Fix treatment as a baseline) on the share of transferred knowledge only for the Non-EDU students. In terms of economic significance, the treatment effect for the Non-EDU students varies between 0.141 (*Bonus*) and 0.160 (*Linear*), so that the introduction of a monetary incentive increases the share of transferred knowledge by at least 14.1 percentage points. Comparing this to the average share of 75% in the *Fix* treatment for Non-EDU subjects shows that this is indeed a sizable increase.

In addition, other potentially confounding variables like age, sex, or school grades have no influence in the full sample. Finally, the only significant variable which deals with the way of teaching is the dummy if instructors have told their pupils that they have the same material at their place. This is somewhat surprising, as a priori the use of the material by the pupils should make it easier for them to understand the topics. However, a possible explanation for the decrease in transferred information might be that the instructor then omits some facts, thinking that the pupil can get them from the material.

3.5 Conclusion and discussion

Financial incentives conditional on the employee's performance are used in many, if not most, professional settings. In the realm of knowledge transfer, however, and especially in the education industry, they are not as widespread. This paper analyzes how different monetary incentives influence the effort and performance of people who are trying to transfer valuable knowledge: Students who want to become teachers are not influenced by different incentive schemes, while other students react in the expected way to monetary incentives: *Linear*, *Bonus*, and *Tournament* pay schemes all improve the ratio of transferred knowledge compared to a fixed wage. Between these three different schemes, there is however no difference. Looking at the time the subjects use to teach instead of the outcome, there is no difference between the different incentive schemes, regardless of the subject pool.

There are two possible explanations for the different results obtained with the two subject pools: People with certain preferences (such as high risk aversion or other-regarding preferences) might self-select into certain fields of study[75]. Dohmen and Falk (2010) show that this is the case for risk attitudes as more risk averse people select themselves into the teaching profession, thereby influencing the effect of certain incentive schemes. However, this can account for only a small part of the results in the paper at hand, as the *Linear* incentive scheme should be influenced by risk attitude less than the *Tournament* or the *Bonus* incentive schemes[76]. In addition, this paper's sample does not exhibit significant differences in risk or trust attitude as measured by a standard survey question in the questionnaire[77]. Furthermore, in most treatments neither trust nor risk attitude are significantly correlated with the ratio of transferred knowledge when tested for each treatment separately[78].

A second and complementary explanation is that (prospective) teachers have a preference for the act of teaching itself. This could mean that such an intrinsic motivation would diminish the effect of an additional extrinsic source of motivation like money[79]. Such a crowding out effect would lead to no effect on teaching quality, matching the observed results. A more likely explanation could be that due to their intrinsic motivation, prospective teachers already give their maximum possible teaching effort, regardless of monetary in-

---

[75]Several studies find for example that students of economic subjects behave differently than students of other fields (Brosig-Koch et al. (2011); Ockenfels and Weimann (1999); Rubinstein (2006)). Brosig et al. (2010) find indication that these differences can be explained by selection rather than education effects

[76]In fact, the only influence of risk attitudes in the *Linear* scheme comes from the uncertainty how the student will perform given a certain teaching level, not from the incentive scheme itself.

[77]Dohmen et al. (2011) show that these survey questions are highly correlated with experimentally validated risk attitudes; Fehr et al. (2002) do the same for the trust question.

[78]This is done with Spearman correlation coefficients. The only significant correlations are between risk attitude and ratio of transferred knowledge in the *Fix* treatment with Non-EDU subjects ($p<0.05$) and in the *Linear* treatment with Non-EDU subjects ($p<0.1$).

[79]See for example Frey and Jegen (2001) or Frey (1997).

centives. This would mean that while introducing an additional monetary incentive might lead to the teachers wanting to increase their effort, this is just not possible. However, this would also mean that in the sample at hand prospective teachers have a lower ceiling in terms of teaching quality than the other students, as they are not able to transfer as much knowledge in the treatments with momentary incentives. An interesting follow-up study would be to look at the long term effect of monetary incentives on knowledge transfer, as the intrinsic motivation for teaching might change over time. This might help to entangle the two possible underlying causes described above.

Chapter 4

Call me if you can - An experimental investigation of information sharing in knowledge networks [80]

4.1 Introduction

The efficiency of research networks and clusters (the latter being characterized by regional links) is dependent on the information flow between the actors involved. Network structures between firms or research institutions have usually developed over a longer time span to achieve a better research output through spillovers between the actors. Policy interventions aim to increase research output by funding cooperation which results in additional network links. While it is well established in the empirical literature that R&D actors increase their output by cooperating in networks, there are ambiguous results on the effects of public policies that aim to promote cluster and network formation (Martin et al., 2011). In particular, not much is known about the effects of different network structures on the efficiency of information flow.

The distribution of information is one aspect of interactive learning in R&D innovation systems (Soete et al., 2010). Presently, about 350 cluster or network organizations exist in Germany alone (Rothgang and Lageman, 2011). These organizations are mainly

[80] See Helbach, C., K. Keldenich, M. Rothgang, and Guanzhong Yang (2012): Call me if you can - An experimental investigation of information sharing in knowledge networks, *Ruhr Economic Papers 332*.

characterized by weak ties between the actors (Granovetter, 1973). However, current research policies influence network structures. One example of such an R&D program is the "Leading-Edge Cluster Competition"[81] in Germany that has changed the network structures substantially, leading to an increase in the concentration of the networks to a few actors (RWI et al., 2011). These observations are the starting point for the analysis at hand, which seeks to answer a general question: If the network structure is taken as given, how does this structure influence the information flow in the network? This question is of practical relevance not only in research cooperations but more generally in cases where network structures have evolved and the question arises whether they should be modified in order to ensure a more efficient information flow.

In network structures, each actor represents a node and the possible communication links represent lines between these nodes. Basic network structures are the star, where one node is linked to all other nodes but no other links exist; or the full network, where all nodes are linked to each other. Most of the existing literature is concerned either with network formation (see Section 4.2 for a brief overview of both the theoretical and experimental literature) or with the effect of network structure in strategic situations, i.e. where a payoff conflict between the actors exists. In contrast, the paper at hand examines a situation without payoff conflicts among the different agents in the network (reflecting a situation where success is only possible when all members achieve a high level of information) and with predetermined network structures. This is done by conducting a laboratory experiment where the participants are assigned to nodes in a network and have to master a task without payoff conflict. Using a laboratory experiment allows the controlled variation of only the variable of interest - namely the network structure - while everything else is kept constant. In addition, agents can be randomly assigned to the different network structures. This way, it is easier to establish causality than in an empirical, non-experimental setting.

Which network is the most successful in terms of information sharing in the real world is not obvious: While many links between nodes of a network allow an efficient sharing of information, they also introduce a coordination problem even if each actor in the network is perfectly rational and the state of the world is common knowledge. Superfluous connec-

[81] The "Leading-Edge Cluster Competition" ("Spitzencluster-Wettbewerb") was launched by the German Ministry of Education and Research in 2007. In this competition, research cooperation in regional partnerships (clusters) is fostered in order to improve innovation (see http://www.bmbf.de/en/10726.php for a detailed description of the competition). While the changes in network structure that take place after participation in the competition are observed, the basic research question remains, what influence these changes have on the efficiency of information flow between the actors involved.

tions might be established or necessary connections might be delayed, leading to a general delay of the spread of relevant information. The paper at hand examines the causes of possible differences in the speed and efficiency of information sharing in different networks. Furthermore, the development of information sharing through several repetitions is studied to see if there are differences between experienced and inexperienced networks.

The paper proceeds as follows. A short literature summary is presented in Section 4.2. The experimental design is described in Section 4.3. In Section 4.4, the hypotheses are stated, while results are presented in Section 4.5. Section 4.6 concludes.

4.2 Literature on networks - knowledge transfer and formation

So far, very few experiments (in the field or in the laboratory) have addressed R&D networks and clusters and the related policy questions even though Sörensen et al. (2010) and Falk and Heckman (2009) argue that experiments are well suited to complement traditional methods like field research. Studies which have already used experiments to analyze R&D-related topics have been conducted by Giebe et al. (2006), who analyze the allocation of R&D subsidies in an experimental setting, and Gächter et al. (2010), who look at knowledge sharing in innovation networks. While the contributions mentioned focus on inefficiencies in selection processes and incentives for knowledge sharing, the paper at hand analyzes network structures and their role for the efficiency of information distribution within research networks.

In the empirical literature, learning and information transfer in R&D and the adoption of new technologies is discussed with respect to diffusion of technologies (Geroski, 2000; Hargreaves Heap and Parikh, 2005) and systems of innovations (Asheim and Coenen, 2005; Edquist, 2005). These studies show that transfer of technological information and knowledge in R&D in general and specifically in R&D networks and clusters follows rather complex patterns. Relevant dimensions of learning in cluster and network structures are interdependencies and interactions between the actors involved (business firms, universities, research institutes and other actors). These interdependencies lie behind network structures, asymmetric knowledge endowments and different resource bases of the actors involved and last but not least patterns of information transfer that exist in R&D clusters and networks.

A question that is closely related to our analysis, the development of network structures, has been addressed by several studies in theoretical models. These studies are relevant for our research question because they can give us hints on the characteristics and practical relevance of different network structures. Jackson and Wolinsky (1996) propose two specific models for the formation of networks and find that - depending on how the costs and benefits of connecting are allocated to the nodes - both the full network and the star network can be efficient and stable structures[82]. Bala and Goyal (2000) use a different model to analyze the genesis of network structures where connections can be initiated by a single node, which then also has to bear all the incurred costs. This allows the modeling of network formation as a non-cooperative game. One notable result of their two-way model (where a connection gives access to the information to both nodes) is that the resulting networks are either empty or a star. In the star network, the central node bears the costs of initializing the connections.

The experimental economic literature on networks can be split into three broad areas[83]: Network formation, coordination networks, and cooperation networks. Again, the first area is the most interesting one for our concerns[84] as it is informative for the decision about which network structures to study in our paper. Falk and Kosfeld (2003) explicitly test the model by Bala and Goyal (2000) and find that the Nash-Equilibrium predictions from this model (namely either an empty or star network) do not hold. Still, in a one-way model the predictions (which are then either empty or circle networks) do hold and the respective networks form. They go on to explain their divergent findings with social preferences which replace the standard, fully selfish preferences in the original model. Furthermore, Goeree et al. (2009) use a laboratory experiment to test the emergence of networks and augment the analysis by introducing different types of agents. They find that the resulting network

[82]In a later model, Bloch and Jackson (2007) enrich this model by allowing players to use transfers (direct and indirect) in their bargaining process.

[83]This classification follows the survey by Kosfeld (2004).

[84]The literature on coordination networks is concerned with the effect of network structures when games have more than one equilibrium, e.g. a risk-dominant and a payoff-dominant one. The results are not clear cut, while Keser et al. (1998) find an influence of network structures on the resulting equilibrium, Boun My et al. (1999) do not (when comparing a circle with a full network). Cooperation networks in contrast use Prisoners' Dilemma or Public Good games to look at the influence of network structure on cooperation. Kirchkamp and Nagel (2000) use a Prisoners' Dilemma game and find less cooperation in a circle network; Carpenter et al. (2010) use Public Good games and find a large influence of network structure on efficiency.

structure strongly depends on the type of agent. The relation of costs and utility associated with a connection is the deciding factor. In particular, stars only develop with one "high utility" agent, not with homogeneous agents or with one "low cost" agent.

Besides economics, social psychology offers two early examples of the experimental analysis of different network structures: Bavelas (1950) and Leavitt (1951). The second paper is especially interesting as it has some similarities to the experiment proposed here. It also analyses the influence of network structures (namely circle, line, Y, and star) on information efficiency. Similar to our design, individuals represent the nodes in the network and can choose to send information along the links in the network. The experimental design is also used by Guetzkow and Simon (1955), who include the full network in the analyzed network structures. Furthermore, their analysis adds a time dimension and finds that the full network is between the circle and the star network in terms of speed. While these studies have some aspects in common with the paper at hand, there are several crucial differences: They do not use monetary incentives, relying instead on the intrinsic motivation of the subjects. Furthermore, the subjects can write free-form messages, allowing mistakes in the information transmission. Finally, only male subjects are used and anonymity is not upheld.

4.3 Experimental design

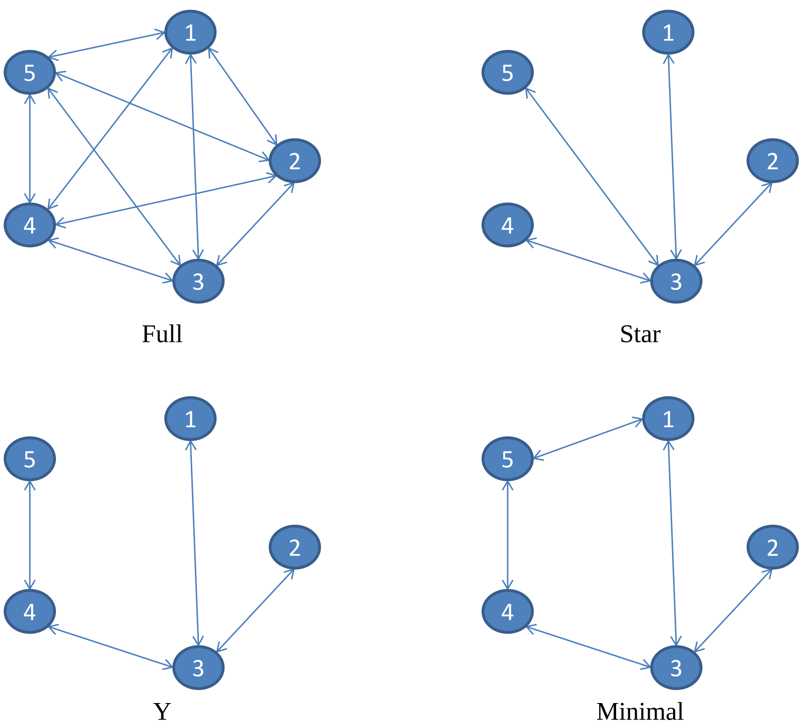
4.3.1 General procedures

The experiment was conducted computer-based and took place at the "Essen laboratory for experimental economics" (elfe) at the University of Duisburg-Essen in July 2011. Participants were recruited via ORSEE (Greiner, 2004) and the attached subject pool. To program the experiment, the software z-Tree (Fischbacher, 2007) was used. In total, 16 sessions with 10 participants each were conducted. The participants were students from the University Duisburg-Essen with an average age of 24.2 years. The sessions lasted at most 90 minutes, the average payoff for the participants was EUR 19.60 with a minimum payoff of EUR 15.70 and a maximum payoff of EUR 23.40.

4.3.2 Treatments

This study aims to investigate networks without payoff conflicts, where every member of the network has the same goal. This goal is defined as the maximum information level for the network, i.e. every member of the network holds all available information. The network structure, i.e. the links along which information exchange is possible, determine several properties of the networks[85]. These properties then allow theoretical predictions which network structure will perform better in terms of information dispersal. From both the theoretical literature on network formation and the actually existing networks in the "Leading-Edge Cluster Competition", three interesting network structures are identified: Full, Star, and Y. A fourth network structure - called Minimal - is used due to its unique properties (see Section 4.4 for a detailed explanation). Figure 4.1 shows these four network structures. In total, 32 independent observations (one observation consists of one network with five subjects each) were gathered, 8 in each treatment.

Figure 4.1: Network structures



[85] A note on terminology: Links denote the channels through which information exchange is generally possible; connections denote those instances when a contact was actually established.

4.3.3 Course of the experiment

The experiment uses five-person networks[86], where each node in the network is an individual subject. A between-subject design is employed where each subject only participates in one treatment. The detailed course of events in the experiment is as follows: Upon entering the laboratory, the subjects are randomly allocated to 10 closed and sound-protected cabins. They receive the instructions (see the Appendix for the translated instructions, all treatments used the same instructions) and have the opportunity to ask questions, which are answered privately by the experimenter. When all subjects have indicated that they have understood the instructions, they have to answer a set of 7 control questions, mainly concerned with the general setup of the experiment and the payoff rules. After all subjects have answered the questions correctly, the experiment itself starts. The experiment consists of 10 rounds with three stages each. For the repetition, a partner matching is employed, i.e. the same network stays together for all 10 rounds. The number of repetitions is chosen such that a convergence of play should be possible for the subjects and learning can be analyzed. Figure 4.2 provides a graphical overview of one round's timing.

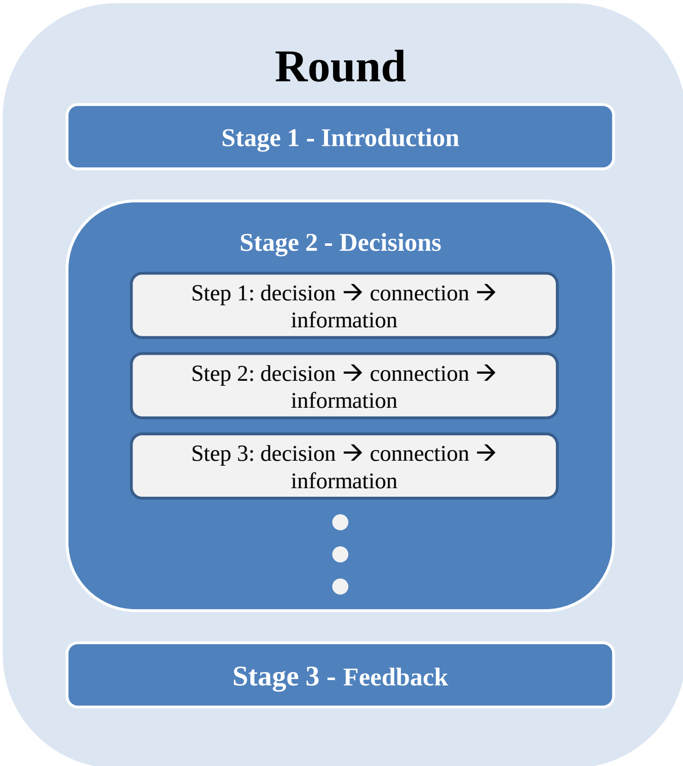
In the first stage, each subject receives a different piece of private information. This information allocation is common knowledge for all subjects. The subjects are informed about the network they are in, and their position in this network. Figure 4.3 shows a screenshot[87].

The second stage is divided into discrete steps. During each step, all subjects can choose one node (or none at all) to connect with. If two subjects select each other, the connection is established and all information both nodes possess is exchanged. If no decisions coincide, no connections are formed. This procedure is repeated until every subject has all available information. From step 2 onwards, the following information is additionally displayed for the subjects: The pieces of information they already possess, the decisions they have made during the current period, and (if applicable) the subject(s) who has (have) tried to contact them in the last step. The design takes great care to avoid any focal points which might influence who is contacted: The network is displayed at a slightly skewed angle, such that

[86] Five-person networks are used because they are the smallest possible networks where reaching full information is not trivial. As this is the first study to analyze the influence of exogenous network structures, this is a natural starting point and allows the clearest identification of the information exchange in the network.

[87] The screenshot is translated from the original German.

Figure 4.2: Experimental design - timing



there is no clear focal point. The different nodes in the network are given the names "lotu", "laja", "leje", "lira", and "lelo" which have been created in such a way as to make ordering them difficult (as opposed to numbers or letters, for example)[88].

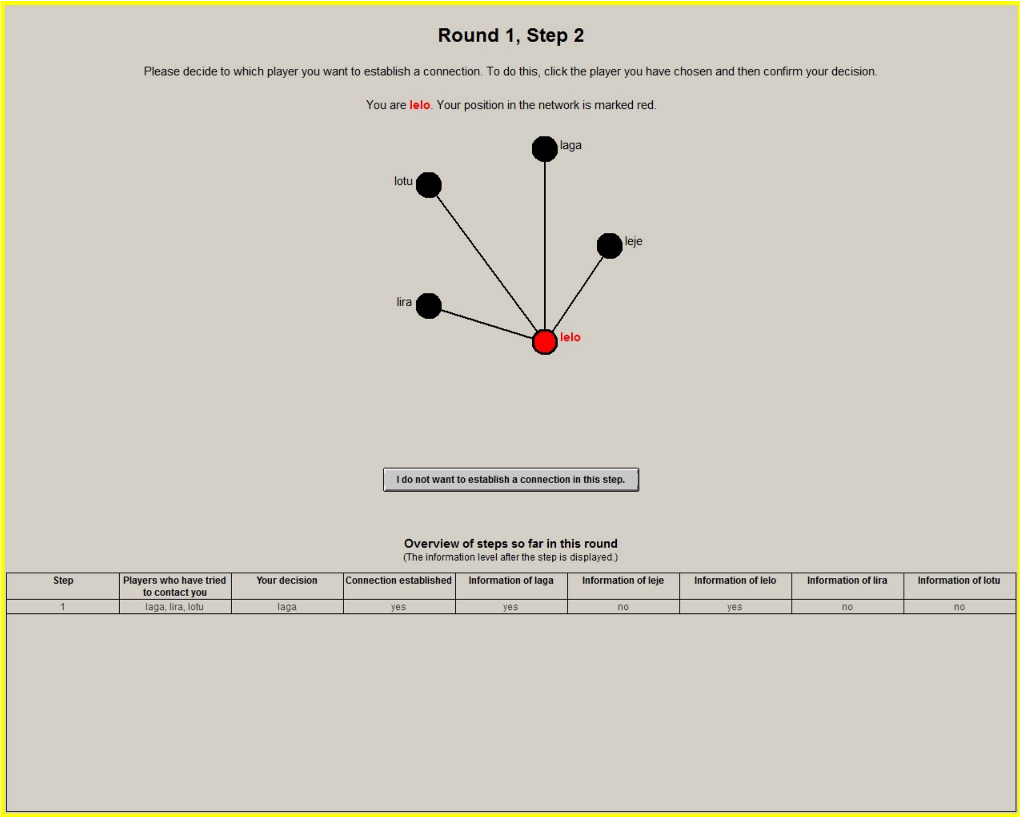
In the third and final stage, the payoffs for the subjects, which are the same for each subject, are computed. Each network starts with EUR 18 per round. From this amount, costs for each step used and each connection established are subtracted where one step costs EUR 0.30 and one connection costs EUR 0.80[89]. Note that connection attempts that do not result in a connection are costless[90]. The network payoff is divided evenly among the members. There is therefore no payoff conflict between the subjects, as the incentives for the group and each individual are perfectly aligned. Throughout the whole experiment,

[88] To test if there really is no "focal point" contained in the name or the geographical position of the nodes, an exact χ^2-test is employed to compare the frequency of choices in the very first decision of the Full network. The relative frequencies are not significantly ($p>0.1$) different from a uniform distribution.

[89] These costs were chosen in a way to achieve a reasonable average payoff (as compared to opportunity costs) for the participants.

[90] The round is ended automatically if the subjects have reached a 0 payoff, i.e. if they needed so many steps and connections that the costs exceed EUR 18. This occurred only once; this group was excluded from the analysis.

Figure 4.3: Screenshot of the decision screen



anonymity is maintained and no communication is possible, except through the described mechanism.

At the end of each round, all own connections are displayed to the members of the network once again. For the final payoff, the payoffs from all 10 rounds are summed up. At the end of the experiment, a questionnaire is filled out by the participants, asking for demographics (like sex, age, study subject), school grades (final cumulative high school GPA[91], last math grade, last German grade), and several questions about their behavior during the experiment (see Section 4.5 for more details of these questions).

4.3.4 Experimental design and characteristics of network and cluster organizations

The experimental design mirrors several characteristics of real world R&D network and cluster organizations. In these organizations, there are many weak ties, i.e. indirect con-

[91] The German "Abiturnote" is used, which is a weighted average of grades received in the last two school years and the main measurement used for university admissions.

tacts or contacts that are not very intensive in respect to number of interactions. Likewise, contacts in the experiment at hand are highly formalized and indirect contacts play a significant role. Of course, different degrees of the strength of the contacts are still possible. These can be influenced by the exogenously given network structure and the endogenously developed routines. Both in the experiment and in existing R&D clusters, information exchange takes place repeatedly, allowing for the emergence of communication patterns as well as improving efficiency through learning.

Furthermore, many cluster and network organizations operate under one common goal or strategy, for which the exchange of information is necessary. This is especially the case if firms and research institutes combine their knowledge to solve precompetitive research topics (Rothgang et al., 2011). Similarly, the incentives for all actors in the experiment are identical. In addition, the relation of the costs for steps and connections in the experiment was chosen to reflect the situation in reality, where a connection - i.e. a visit to a cooperation partner in a research project - is more costly than a unit of time per se. Finally, three of the network structures used in the experiment (Star, Y, and Full) are - as already mentioned - often found in cluster organizations, e.g. in the "Leading-Edge Cluster Competition" in Germany.

4.4 Hypotheses and benchmarks

4.4.1 General network characteristics

First, there are some generally descriptive properties of the networks. These are the total number of links in the network, the average number of links per network node, and the variance of the number of links per node; see Table 4.1.

Furthermore, the following properties can be defined: The minimal number of connections needed to give every node all of the information, the minimal number of steps needed to give every node all of the information, and the number of nodes that on average do not communicate through the optimal path to complete information even though they could. This last property serves as a measure of how much coordination is needed in the network to reach the optimal path. In the Full network, for example, subjects only need 6 connections in 4 steps to establish full information, making a maximum profit of EUR 12 for the group possible[92]. However, on average 1.5 nodes do not communicate along an optimal path indicating a notable coordination problem.

Table 4.1: Network properties

Properties	**Full**	**Star**	**Y**	**Minimal**
Number of links	10	4	4	5
Average number of links	4	1.6	1.6	1.8
Variance of links	0	1.8	1.44	0.2
Best possible result				
Connections needed	6	7	7	6
Steps needed	4	7	5	4
Nodes not communicating	1.5	0	0.8	1.25
Random				
Average connections	11.56	15.66	18.27	15.24
Variance of connections	9.18	28.72	48.45	24.75
Average steps	28.98	39.15	36.54	29.66
Variance of steps	111.71	237.08	238.41	133.01

Source: Own calculations, simulation results after 100,000 runs.

As achieving the optimal possible outcome is very demanding for the ability of the individual nodes to form the correct connections (especially in the networks where there is a coordination problem), it is also useful to look at a lower benchmark. To this end, purely random behavior by the nodes is simulated. From this simulation, the following network properties are derived (after 100.000 repetitions): The average number and the variance of connections needed to achieve full information and the average number and the variance of steps needed to achieve full information. Table 4.1 gives all these properties for the networks which are considered for this study.

Note that for the networks without any coordination problem (Star and Y), the best achievable outcomes can be expected when players are fully rational. In the Minimal network, where there is only a slight coordination problem, optimal play would result in 4.5 steps and 6 connections[93]. The severity of the coordination problem in the Full network, however, will make it very unlikely that the best possible result is achieved[94].

[92]See Figures A - D in the Appendix for concrete examples of an optimal path for every network.

[93]See the Appendix for a detailed derivation. If the additional assumption is made that a group is able to repeat the optimal result if it has reached it once, the steps needed decrease to 4.09 per round if all ten rounds are examined.

[94]It is furthermore very difficult to assess rationality in the Full network from simply observing decisions. As players usually do not know which information the other players have, in most cases it is impossible for them to exclude one potential node completely. Instead, most decisions they can make are rational with a positive probability, depending on the expected allocation of information.

4.4.2 Hypotheses on treatment effects

As mentioned in Section 4.1, a conflict exists between more possibilities to exchange information and possible coordination problems. This conflict is smallest in the Minimal network, as this structure is chosen so that all information can be exchanged as quickly as possible with the smallest number of links. The coordination problem can be expected to have the largest negative influence at the beginning of the experiment, as the players have not yet gained any experience. This leads an ordering of the networks according to the severity of the coordination problem. In particular, the Full network can be expected to perform worst.

Assuming that the coordination problem gets less influential in the course of the 10 rounds, the Full network should improve. At the same time, the Star network should perform relatively worse due to its efficiency constraints. Combining these arguments and looking at the whole experiment the following hypothesis is derived:

H1: *"Among the four different networks, the order of networks in terms of average profit over all rounds will be as follows: Minimal = Y > Full = Star."*

This hypothesis might also serve as an indicator of how severe the coordination problem really is. If e.g. the Star network is more successful than one of the others, the benefit given by the higher number of links is not high enough to overcome the coordination problem.

4.4.3 Learning and convergence

As the game is repeated ten times with the same group, one can observe whether learning occurs i.e. if the participants achieve a higher profit in the later rounds. The game itself is not trivial for the participants to master, therefore one might expect that they do not achieve the best possible result in the very first rounds, leading to the second hypothesis:

H2: *"In all four networks, the participants will achieve a higher profit in the last five rounds than in the first five rounds."*

Another way to look at the development of results over time is to analyze whether a convergence of play can be observed, i.e. if the changes in behavior from round to round become less or even disappear. To do this, several possible strategies can be used. First, the number of consecutive rounds - counting from the last round - with the same result in terms of connections and periods can be counted. The higher this number, the earlier an equilibrium (as in a situation with resistance to change) is reached. If this number is 1, the group has not converged to a certain pattern of behavior at all. This serves therefore as an aggregate measure of convergence in the groups.

Second, the individual decisions can be analyzed by looking at the share of identical individual decisions from one round to the next in one single step. If, e.g., one subject chooses the same connection in the first step in every round, her behavior is very stable. Looking at this share on a group level gives a good idea how far this group has converged to one path of connections.

Finally, one can again count consecutive runs from the end in which a group behaved identically in a certain step. E.g., if every group member makes the same decision in step two in the last five rounds, this might serve as an indicator of a relatively high convergence in that group.

4.5 Results

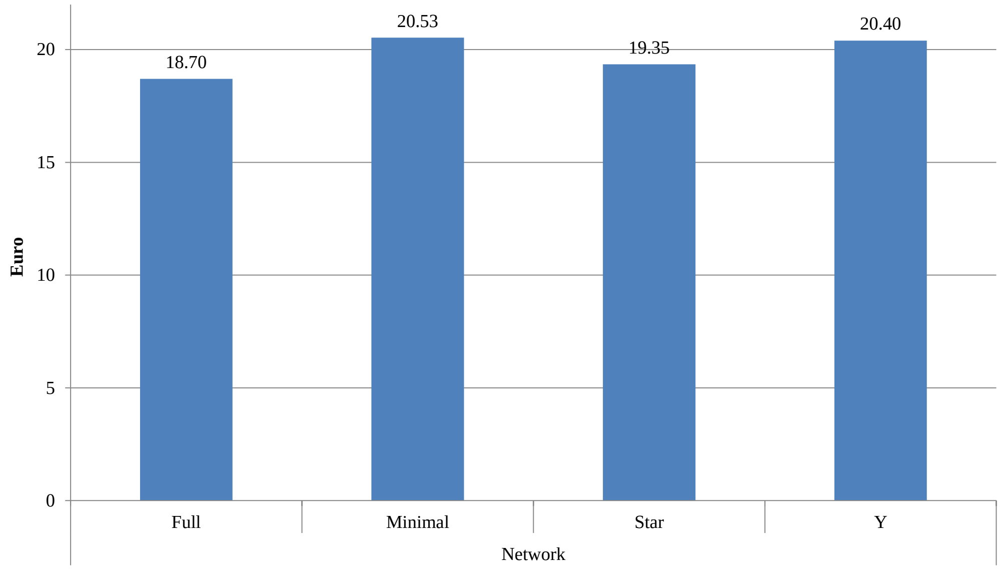
4.5.1 Treatment effects

First, the differences in behavior and outcome between the treatments are analyzed. Figure 4.4 shows the average profit the different network structures achieved in all ten rounds. At first glance, the Minimal and Y networks manage to reach a higher profit than the other two structures, especially in the later rounds. The total profit (over all 10 rounds) for the networks are EUR 18.70 for the Full, EUR 19.35 for the Star, EUR 20.40 for the Y, and EUR 20.53 for the Minimal network. A Kruskal-Wallis test confirms that there is a significant difference in profit between the treatments (p=0.026)[95].

Performing then pairwise Mann-Whitney-U tests (again for average profit over all 10 rounds) reveals that Minimal and Y are both significantly higher than Full and Star; while there is no difference between Minimal and Y on the one hand and Full and Star

[95]If not indicated otherwise, all tests are two-sided and exact.

Figure 4.4: Average total profit by treatment



on the other hand[96]. Looking at these results, hypothesis H1 cannot be rejected: While the Minimal network is indeed the most efficient one, it is no different to the Y network. Similarly, Full and Star perform similarly, showing that both too many links and very restricted communication channels are harmful.

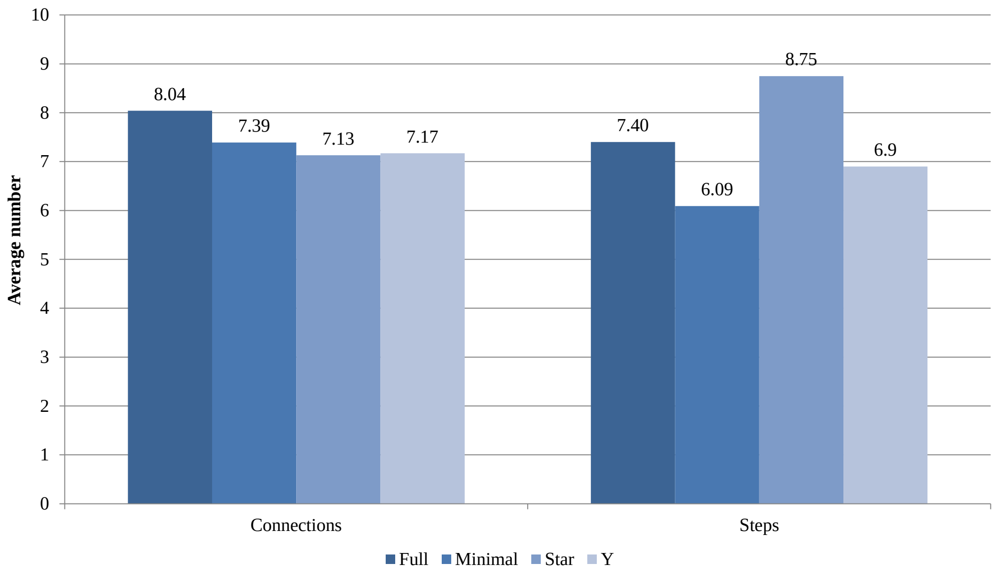
As profit is just a linear combination of steps and connections needed per round, the question remains as to which of the two variables drives this result. Figure 4.5 shows the average number of steps and connections needed to achieve full information in each round. The numbers for connections are 8.04 for the Full, 7.39 for the Minimal, 7.17 for the Y, and 7.13 for the Star network. Here, the one clear difference seems to be that groups playing in the full network seem to need more connections than the groups playing in the other networks. Again using pairwise Mann-Whitney-U tests, this is confirmed, as the Full network needs significantly more connections on average than the other three networks[97].

The second driving factor for profit are the steps needed. The picture is somewhat different compared to the connections: The Star network is now the one needing more steps than the others; the numbers are 8.75 for the Star, 7.40 for the Full, 6.90 for the Y,

[96]The p-values for the pairwise comparisons are as follows: Star vs. Y: 0.015; Star vs. Minimal: 0.040; Full vs. Y: 0.038; Full vs. Minimal: 0.072; all other comparisons are insignificant with p>0.1.

[97]P-Values are as follows: Full vs. Star: 0.000; Full vs. Y: 0.001, Full vs. Minimal: 0.058; all other comparisons are insignificant with p>0.1.

Figure 4.5: Average total connections and steps by treatment



and 6.09 for the Minimal network[98]. Looking back at the differences in profit, the lower profit by the Full network is driven by too many connections, while the lower profit of the Star network is driven by too many steps. The most likely explanation in case of the Full network is the coordination problem. Every node in this network structure has four links, making it hard to establish any connection in the first place. This amplifies the problem for the individual nodes to pick useful connections, driving up the total number of connections. In the case of the Star network, the many steps at the beginning are likely due to getting familiar with the decision situation (this is of course similar for all networks). The difference, however, stems from the natural limit in steps: in the Star network, the smallest number of steps which is possible to achieve is seven, as opposed to four in the Full and Minimal networks and five in the Y network. This means that even when all networks move towards these limits in the rounds, the Star networks will still need more steps on average.

After the experiment, subjects filled out a questionnaire with the following information[99]: demographic information, school grades, questions concerning their strategies in the experiment, and questions concerning their satisfaction with the experiment[100]. How-

[98]This is again confirmed by using pairwise Mann-Whitney-U tests (The p-values are as follows: Star vs. Full: 0.099; Star vs. Y: 0.007; Star vs. Minimal: 0.000; all other comparisons are insignificant with $p>0.1$.)

[99]See Table A.4.1 in the Appendix for a complete list of questions.

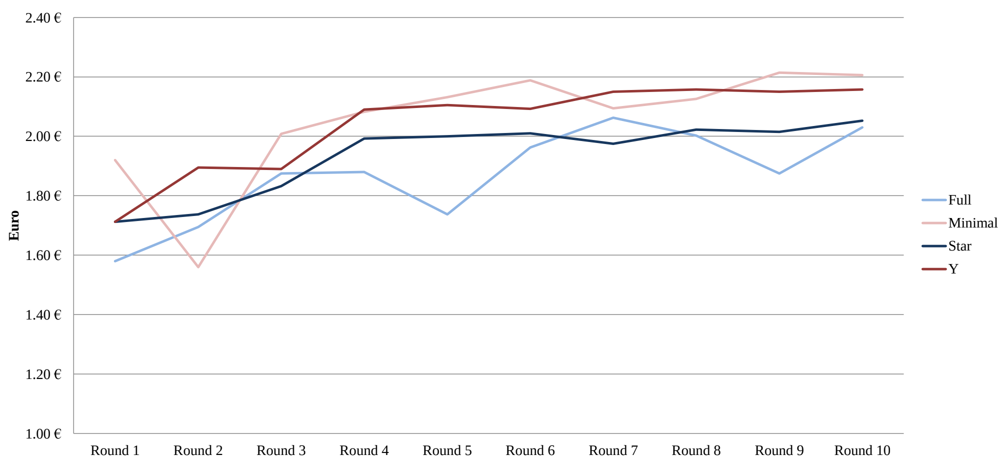
[100]There is no difference according to a Kruska-Wallis test between treatments in average age, average number of semesters, average math grade (scaled to reflect different kinds of course), average German

ever, of special interest from the questionnaire variables is the subjective happiness with the process, quantified by asking for the number of rounds in which the subjects were satisfied with the results. This is especially relevant from a practical point of view. Thinking back to the evaluation of R&D clusters, the perception of the participants might be just as important as the actual results of the cooperation. Using Spearman correlation coefficients for every network separately reveals that while the number of connections is not correlated to the overall satisfaction ($p>0.1$ for all networks), the number of steps is negatively correlated in all networks ($p<0.1$). This is somewhat surprising, as the payoff is per definition negatively correlated to both connections and steps. Indeed, connections are even more costly than steps. A possible explanation is that subjects value the (successful) connections per se and see them as progress towards the common goal of information exchange. In addition, they may not be able to identify unnecessary connections. A second possible explanation is that participants not only take the monetary costs of steps into account, but are also generally impatient and attach costs to waiting until the round is finished.

4.5.2 Treatment effects over time

The next step in the analysis is to see whether the treatment effects described in the previous section change in the course of the 10 rounds. Figure 4.6 shows the average profit of the different networks in all ten rounds.

Figure 4.6: Average profit by treatment, all 10 rounds



grade (scaled to reflect different kinds of course), number of males in the group, and number of economists in the group.

To see if there is a difference in the networks, the results from the first five rounds are compared to the results from the last five rounds. In the first five rounds, hardly any clear pattern can be seen in the average profit, testing for pairwise differences shows that only one difference - Minimal achieving a higher profit than Full - is significant (comparing the average profit of the first five rounds, p=0.044). The pattern becomes clearer when looking at the last five rounds, however. The average profit of the last five rounds is significantly higher for Y and Minimal than for Full and Star[101]. The overall pattern of treatment differences is therefore driven by the results in the last five rounds.

Summarizing treatment effects, the Minimal and the Y networks are better in terms of profit than the Star and Full networks. This difference is due to the high number of connections in the Full network and the high number of steps in the Star network. These results are mainly driven by the behavior in the second half of the experiment.

4.5.3 Learning and convergence

The structure of the experiment also enables the dynamic process the different networks go through to be analyzed. This is also interesting with regard to practical applications of the research question. R&D networks, for example, in most cases exist for a longer time period. Usually, the interaction among the different agents is repeated with the same information structure. To see whether groups improve their performance during the course of the experiment, the first five rounds are compared to the last five rounds for each treatment. Looking at Figures 4.6, 4.7, and 4.8, the decisions on average seem to improve, with the groups needing fewer steps and connections and therefore achieving a higher profit in later rounds.

This is confirmed by a Wilcoxon-Signed-Rank testing for differences in the steps needed to achieve full information in each network (All p-values are<0.02). For the other target variables profit and connections, only the Y network improves significantly when comparing the first half of the experiment to the second half (p=0.031). Hypothesis H2 therefore can be partially rejected: subjects do indeed need fewer steps in later rounds, but only one network structure (Y) also shows significant differences in profit and connections.

[101] The p-values are as follows: Y vs. Full: 0.049; Y vs. Star: 0.000; Minimal vs. Full: 0.099; Minimal vs. Star: 0.003; all other comparisons are insignificant with p>0.1.

Figure 4.7: Average connections by treatment, all 10 rounds

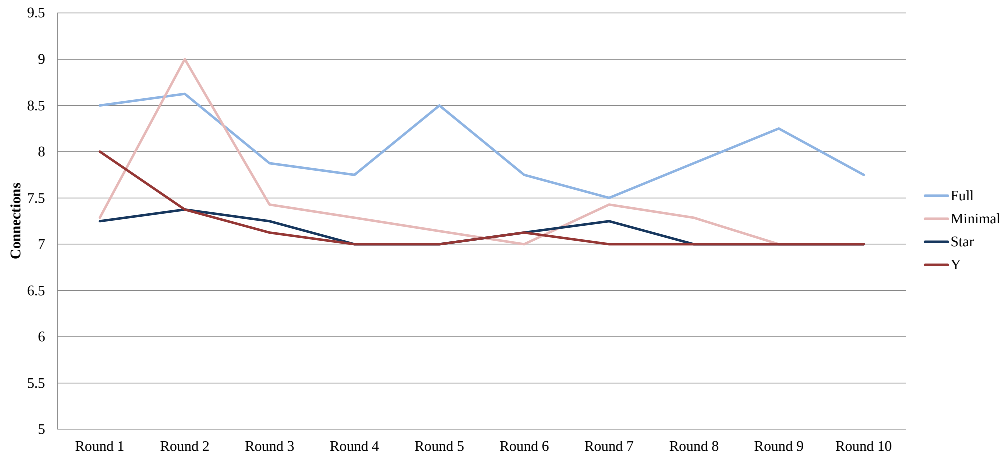
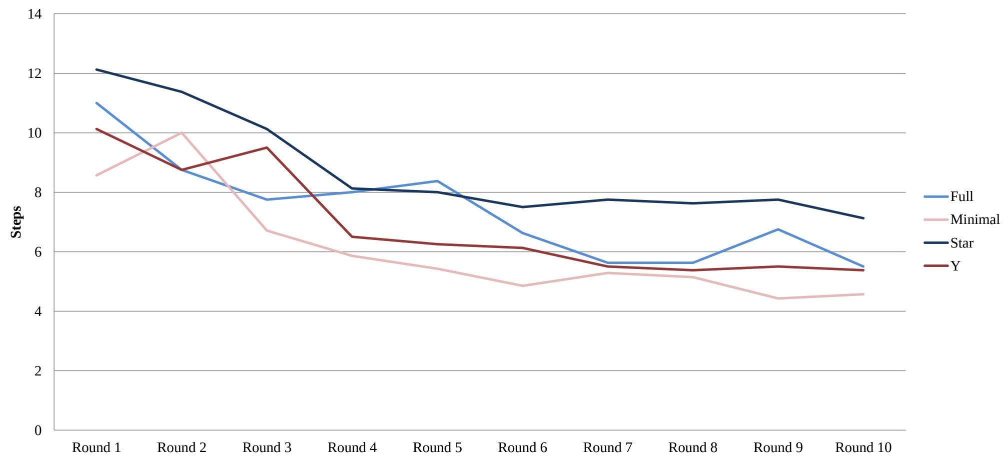


Figure 4.8: Average steps by treatment, all 10 rounds



Section 4.4.2 describes three different ways to look at possible convergence over time in the networks. Using the first method, one looks at the number of consecutive rounds (counted from the end) with the identical result. On average, these are 2.25 in the Star network, 4.5 in the Y network, 2.25 in the Full network, and 3.43 in the Minimal network. Testing these results with pairwise Mann-Whitney-U tests, groups in the Y network converge significantly faster than those in the Star (p=0.027) and Full (p=0.038) network, but not than those in the Minimal network. Further significant differences cannot be found.

The second method uses the average ratios of all decisions in step 1 to 4 which were identical in the same step of consecutive rounds. Comparing Star, Y, and Minimal, no statistically significant differences can be found. Only the Full network shows less convergence compared to the other network structures: compared to the Star and Y network in steps 3, 4, and the average of the first four steps and compared to the Minimal network only in step 4.

Finally, the last method investigates in how many consecutive runs - counted from the end - a group's decision is exactly identical in step 1 to step 4. The differences in this number between the network structures are not significantly different (using a Kruskal-Wallis test). However, this number increases from the first five rounds to the last five rounds in all networks, indicating that all networks do indeed move into the direction of equilibrium behavior in the course of the 10 rounds.

Summarizing the results on the dynamics from the experiment, all networks need fewer steps in the later rounds, while only the Y network also improves in profit and connections. Regarding convergence, the results are not as clear cut and depend on the method used. However, it seems that most do indeed converge to some equilibrium behavior in the later rounds.

4.5.4 Individual decision making

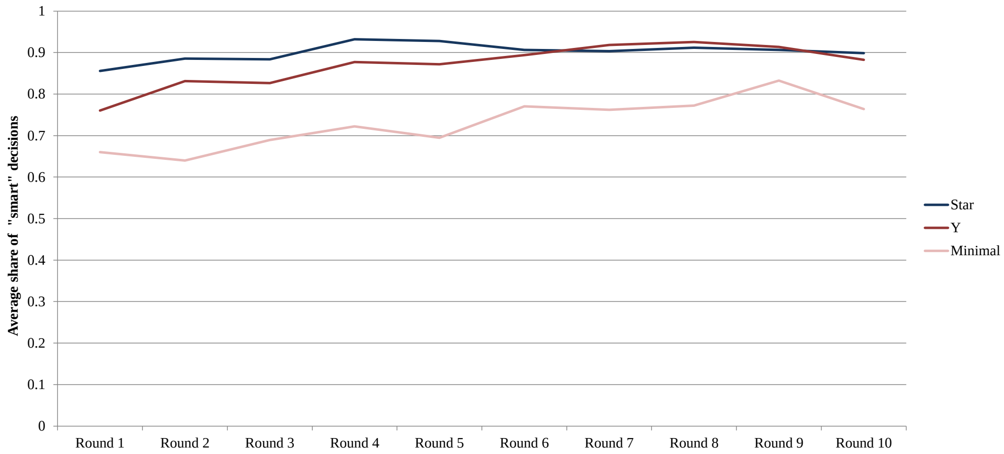
In addition to the aggregate behavior analyzed in the preceding sections, one can also look at individual decision making in the networks. To this end, the networks where there is no (or only a very small) coordination problem - i.e. Star, Y, and Minimal - are analyzed separately from the Full network which suffers from a large coordination problem.

In the first three networks mentioned, it is possible to classify individual decisions as "smart" (i.e. maximizing the expected payoff) as the coordination problem is not so severe[102]. Figure 4.9 shows the share of "smart" decisions by all subjects in the three different networks.

Reflecting the slight difference in their coordination problem, subjects in the Y and Star networks make the fewest mistakes and consistently play "smart", while subjects in the Minimal network make "smart" decisions somewhat less frequently. Still, the overall share of "smart" play is pretty high, with a slight trend towards better decision-making in the later rounds. This reflects the aggregate results of higher average payoffs in the later rounds as subjects learn to avoid mistakes in the later rounds and therefore make better

[102]To avoid classification errors, the coding was done by two persons separately. Figure A.4.5 in the Appendix shows the share of identical classifications between the two coders.

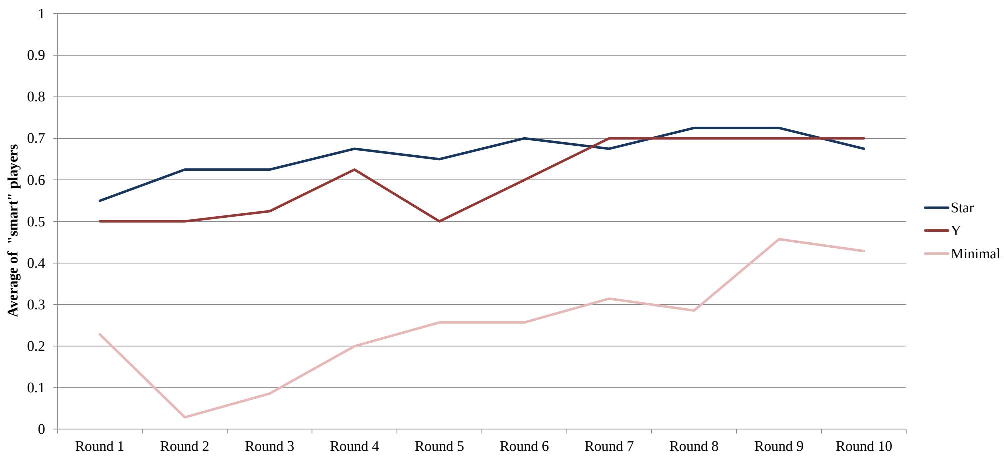
Figure 4.9: Average share of “smart” decisions



decisions. The scope for this learning is largest in the Minimal network, resulting in a clearer upward trend.

A second way to look at the individual decisions is to classify the subjects into “smart” and “non-smart” players. A subject is classified as “smart” if at least 90% of her decisions are “smart”. Using this classification makes it possible to see whether the results described above are driven by single subjects or whether all subjects behave the same. Figure 4.10 shows the share of “smart” players in the different networks.

Figure 4.10: Average share of “smart” players



Again, Y and Star are different from Minimal, with a consistently higher share of “smart” players. All networks again exhibit an upward trend, pointing again to the fact that subjects learn to avoid mistakes in the course of the 10 rounds. Another interesting observation is that for the Star and Y networks, there is hardly any improvement in the last four rounds. One possible explanation would be that all subjects capable of understanding the situation have done so by round six and are therefore classified as "smart", while the rest of the players keep on making mistakes even through repetition of the game.

The Full network suffers from a severe coordination problem, so it is hardly surprising that the best achievable outcome is not realized. There are mainly two pieces of information which might serve as a guide for the players in the decision situation: The connection attempts of the other players in the preceding step and the player(s) whose information they are still missing. To see if players actually use these information (and which is more important), two simple heuristics are investigated:

1. Connect with somebody whose information is still missing. If you already have all the information, choose randomly (called "No info").
2. Connect with somebody who has tried to connect with you in the preceding period. If nobody tried to connect with you, choose randomly (called "Call back")[103].

For both of these heuristics, two rules additionally hold: subjects never call themselves (which is the equivalent of not trying to establish any connections) and never try to contact the subject with whom they have just established a connection. Table 4.2 shows the simulated results from these heuristics and as a comparison the actual results from the first two rounds[104].

Table 4.2: Heuristics in the Full network

Heuristics	**Connections**	**Steps**
No Info	9.10	12.90
Call Back	10.40	18.23
Actual results		
Round 1	8.50	11.00
Round 2	8.63	8.75

Source: Own calculations, simulation results after 100,000 runs.

In addition to this aggregate view of behavior in the Full network, all individual decisions were classified as either compatible with the two heuristics or not. This results in a share of 0.66 of decisions that are compatible with the "Call back" heuristics and a share of 0.86 decisions that are compatible with the "No info" heuristic in the first round. Taking

[103]For this heuristic, an additional rule has to be followed: If you have not established a connection in the preceding step and somebody tried to connect with you, choose randomly between calling him back and calling somebody else. If this rule is not added, the possibility exists that no connections are established, resulting in a zero profit.

[104]The first two rounds are used as the results from later rounds are very path-dependent and therefore not suited to evaluate heuristics.

these two approaches together, it seems that the "No info" heuristic is able to explain the data pretty well, certainly better than the "Call back" heuristic. Subjects thus seem to concentrate on collecting all available data, not on establishing a connection per se. However, the actual results from the first rounds are still better than the results simulated with the heuristics, especially concerning the steps needed. One explanation is that the subjects find some kind of coordination device not covered in the heuristic. As mentioned before, the distribution of connection attempts between the different nodes in the Full network is no different ($p>0.1$) from an equal distribution. Refining this analysis, however, reveals that subjects try to connect with one of their geographic neighbors[105] significantly more often ($p<0.05$) than with the other two subjects. Therefore, the geographic location might serve as an additional coordination device, improving the heuristic and making fewer steps possible[106].

As a short summary, the coordination problem plays a big role (as expected) for the individual decisions. In the network structures without a coordination problem, the subjects are able to play close to the optimal outcome, especially in later rounds. For the Minimal network with its small coordination problem, it gets harder for the participants to play this way. The large coordination problem in the Full network makes it necessary for the participants to rely on heuristics to achieve some coordination. They mainly use their own information status as a guide who to contact next.

4.6 Conclusion

The experimental results suggest that the network structure strongly influences the speed and efficiency of information exchange. As expected, increasing the number of possible links in a network has positive (due to more possibilities of information exchange) and negative effects (due to coordination problems). In our experiment, the most efficient network structures are those that find a compromise between these two conflicting aspects, namely the Minimal and the Y network.

Consequently, the results of this paper suggest that cluster and network policies should not pursue the goal to maximize the number of links between the actors involved. In

[105] By geographic neighbor, the two nodes closest on the screen - i.e. along a virtual circle - are denoted.

[106] In the data, this rationale can unfortunately not be distinguished from the heuristic without it, as one still has to assume that every participant is chosen with a positive likelihood.

cluster and network structures with only a few links, impulses for additional cooperations will probably increase the efficiency of information sharing. Additional network links in networks that feature a lot of links from the beginning, however, might result in making information sharing and coordination less efficient. Instead, behavior of participants in the Full network suggests the usage of a simple heuristic: participants try to connect with group members whose information they are missing. On the one hand, this helps to achieve relatively favorable outcomes; on the other hand, it prevents the group from reaching efficiency. Furthermore, participants' satisfaction is positively correlated with the number of connections they establish. This could be one reason for superfluous connections and thus limit learning in the way observed. Besides, policies which use a competition to reward promising R&D clusters might use the structure of information exchange as an additional evaluation criterion. Turning to the organizations themselves, our results suggest that the management of the information flow should be an important aspect of the work of a cluster management, especially if cooperation is planned for a longer term. While such measures may incur costs, clusters should accept that coordination problems are real and harmful.

Learning takes place in all network structures, as groups are able to decrease the steps needed over the course of the ten rounds. However, only groups in the Y treatment manage to improve profit and decrease connections significantly. It is furthermore interesting that the differences between the networks only come to light in the second half of the experiment. One must therefore distinguish between experienced and inexperienced networks. In the former, the network structure is a decisive factor in the efficiency of information exchange. Thus, the analysis conducted here is more relevant to longer existing networks and older cluster initiatives.

Despite the prevalence of learning, a consistent pattern of convergence is not observed. This result may be driven by the restricted and ex ante known number of rounds. It might be interesting for future research to relax this restriction. Generally, the experimental design simplifies the real life situation. For the analysis of R&D networks, situations with imperfect or tacit information diffusion, as well as variations of incentives, information distribution, and group size are natural extensions that should be analyzed.

Chapter 5

The more you know? Consumption behavior and the communication of economic information [107]

5.1 Introduction

Saving money for the future is among the most central economic decisions private households have to make. Therefore, the decision how to split available income between saving and consumption has always been of major interest in economic research. Accordingly, there is a large literature both theoretical and empirical trying to explain the savings behavior of economic agents (see e.g. Schunk, 2009 for an introduction). Given the complexity of the saving decision, research focuses on disentangling the importance of certain factors that influence saving behavior. These include a great number of economic, psychological, sociological, and institutional factors. Browning and Lusardi (1996) provide a very comprehensive survey on consumption and saving. They list a total of nine possible motives why people save, including the precautionary, the life-cycle, the intertemporal substitution, and the enterprise motives. Disentangling the importance of one single motive is extremely difficult. However, it is agreed that the uncertainty surrounding income (and expenditures) is one of the main determinants.

[107] See Brosig-Koch, J., K. Keldenich (2012): The more you know? Consumption behavior and the communication of economic information *Ruhr Economic Papers 387.*

The early theoretical work on consumption and savings faced the problem that a realistic representation of the savings decision (based on realistic preferences and/or stochastic income, among others) is very difficult to solve analytically. One way to deal with this is to introduce a more tractable utility function. Following this approach and using quadratic utility functions, the Certainty Equivalent model was developed (Deaton, 1992). However, the results obtained with this model are not able to explain several empirical observations regarding savings (see for example Carroll, 2001). Thus, the second way taken to model consumption behavior was to solve the complicated model numerically and not analytically. This approach resulted in the Buffer Stock Savings Model (Deaton, 1991; Zeldes, 1989) which predicts that consumers should build up a "buffer stock" of savings to prepare themselves for future negative income shocks. This model forms the theoretical base for the paper at hand.

Besides empirical research (see for example Carroll, 1997), the latter model also motivated a number of laboratory experiments. Such experiments allow a high degree of control over the decision environment and, thus, are particularly useful for testing the predictions of this model. More specifically, uncertainty can more easily be disentangled from other determinants of saving behavior. In a pioneering experiment, Hey and Dardanoni (1987) found that subjects were unable to consume and save optimally as predicted by economic theory, but that the comparative static predictions of the theory hold nevertheless. This result was confirmed by Carbone and Hey (2004) who also report significant over-reactions to changes in risk. The experimental analysis by Brown et al. (2009) introduces the possibility of learning how to save optimally. They find that both "social learning" (by seeing the decisions of other participants) and individual learning (by repeating the sequence of decisions seven times) improve the quality of decision making.

This paper builds on the experimental design by Brown et al. (2009) and aims to investigate what kinds of information exactly influence the agents' decision. What should people know about the random process affecting future incomes in order to make optimal decisions? Do other people's beliefs about the future income already affect behavior positively - or do subjects need to be informed about others' choice to improve their consumption decisions as found by Brown et al. (2009)? Furthermore, since the future income is affected by various economic events (like unemployment or changes in tax policy), subject's connotation of these events might also affect their consumption. Accordingly, we test whether the framing of the source of the income shock has an influence on behavior.

5.2 Model

The model used for the paper at hand is a Buffer Stock Saving Model and largely follows Brown et al. (2009), who in turn base their parameters on empirical evidence reported by Carroll et al. (2000). Consumers live for 30 periods. In each of these periods, subjects earn a stochastic income and have to decide how to allocate their cash-on-hand between saving and consumption. The utility function is a CRRA (constant relative risk aversion) one and incorporates a habit stock into the calculation:

$$U(C_S, H_{S-1}) = \frac{(\frac{C_S}{H_{S-1}^{\gamma}})^{1-\rho}}{1-\rho} \tag{5.1}$$

The following table gives the variables used in the model:

X_s	- Total cash / funds available in Period s (denoted “cash-on-hand”)
S_s	- Savings in period s (Share of X_s not used for consumption)
C_s	- Consumption in period X_s
R	- Gross interest rate in each period
H_{s-1}	- Habit stock from period $s-1$
$U(C_S, H_{S-1})$	- Utility in period s
Y_s	- Actual income in period s
P_s	- Fixed income in period s
g_s	- Growth rate of fixed income in period s
G_s	- $1+g_s$, indicates the development of the fixed income: $P_{s+1} = P_s G_{s+1}$
η_s	- Stochastic income shock in period s
ρ	- Coefficient of relative risk aversion
γ	- Parameter to indicate the importance of the habit stock variable

The inclusion of a habit stock (H_t) into the calculation of course makes early consumption less attractive as it decreases the utility of future consumption. As a consequence, early saving becomes more attractive even above the need to insure oneself against future negative income shocks. The consumer maximizes his expected lifetime utility, leading to the following maximization problem:

$$\max E_t[\sum_{s=t}^{T} U(C_s, H_{s-1})] \tag{5.2}$$

In order to keep the experimental task manageable, the time preference factor is assumed to be 1 and is therefore omitted. The habit stock develops according to the subjects' consumption: $H_t = (1-\sigma)H_{t-1} + C_t$, where σ is the habit stock's depreciation rate. This rate indicates how persistent the habit stock is. The higher σ, the slower the habit stock decreases back to zero and the more important is it to incorporate it into one's calculation. In this model, no borrowing is allowed (i.e. $S_t \geq 0$). To simplify computation, it is convenient to divide all variables by fixed income P_s, thereby eliminating the fixed income variable. Lower case variables denote variables where this division has taken place. A recursive specification of current and future utility can now be written as a function of only two variables, namely h_{t-1} and x_t. The optimal value function is then

$$\overline{V}_t(x_t, h_{t-1}) = maxu(c_t, h_{t-1}) + E_t[\overline{V}_{t+1}(\frac{R}{G}[x_t - c_t] + \eta_{t+1}, \frac{\sigma}{G}h_{t+1} + \frac{1}{G}c_t)] \tag{5.3}$$

Subject to constraints

$$s_t = x_t - c_t \tag{5.4}$$

$$x_{t+1} = \frac{R}{G}s_t + \eta_{t+1} \tag{5.5}$$

$$h_t = \frac{\sigma}{G}h_{t+1} + \frac{1}{G}c_t \tag{5.6}$$

Note that the first constraint implies that there is no borrowing; subjects can only spend their available cash-on-hand which is dependent on past savings and current income.

Following the parameter specification used by Brown et al. (2009), the utility function is adapted to include some scaling parameters, and parameters were chosen such that $\rho = 3$ and $\gamma = 0.6$. The following utility function was therefore used in the experiment:

$$U(C_t, H_{t-1}) = 40 + 750\frac{(\frac{C_t+2.7}{H_{t-1}^{0.6}})^{1-3}}{1-3} \tag{5.7}$$

Furthermore, the stochastic income shock has to be specified. The income shock η is chosen to be lognormally distributed with a standard deviation of $\sigma^2 = 1$: $log\eta \backsim N(-0.5, 1)$. This results in a mean income shock of $E[\eta] = 1$. The growth rate of the fixed income is set to be constant at 5% per period. The habit stock's depreciation rate σ is 0.3 which signifies a slow decrease of the habit stock over time. The starting habit H_0 is set to 10.

5.3 Experimental design

At the beginning of the experiment, subjects receive detailed instructions and have to answer eight comprehension questions about the decision task.[108] In total, subjects played thirty periods. The first period only starts after all subjects have correctly answered all questions. During the experiment, possible questions asked by participants are always answered in private by the same experimenter. In the description of the decision task, technical terms from economics are avoided (thus the income shock is called "adjustment factor" and the habit stock "lifestyle index"). The money in the experiment is given in "Experimental currency" (EW from the German "Experimentalwährung") units. The amount of EW they choose to spend in a period is converted to actual Eurocents. Participants are given a large table in the instructions showing possible results from different spending decisions. In addition, they also receive the formula and a table indicating the development of the habit stock. In each period, participants are informed about the current income, their available cash-on-hand, and their habit stock. They further receive the information that their income is subject to a random adjustment factor and that this adjustment factor is drawn independently in each period. To illustrate this, subjects are given three example draws of the income shock over all 30 periods. To test what a certain amount of spending would mean in terms of Euro payoff, subjects could use a Euro calculator which was implemented in the decision screen. The use of this calculator is recorded for future analysis. Before each decision, subjects have to state a belief about this period's adjustment factor. Figure 5.1 shows a screenshot of the decision screen.

In total, we varied three treatment dimensions: The specifics of information subjects receive about the random process determining their income (Info treatments), the availability of other subjects' beliefs about the current period's adjustment factor (Belief treatment),

[108] See the Appendix for a full set of instructions.

and the framing of the source of income shocks, i.e., the framing of the random adjustment factor (Frame treatments).

We implemented four Info treatments in addition to the baseline treatment (Base) described above. In treatment LowInfo, subjects do not receive any example draws of the income shock. Treatments PointInfo and IntervallInfo add a point estimate and an interval estimate of the random variable to the three example draws, respectively. The HighInfo treatment implements the most detailed information concerning the income shock and adds the density function of the random variable, a statement that the income shock is always larger than 0 but will be smaller than 1 most of the time, and the 90% confidence interval for the random variable.

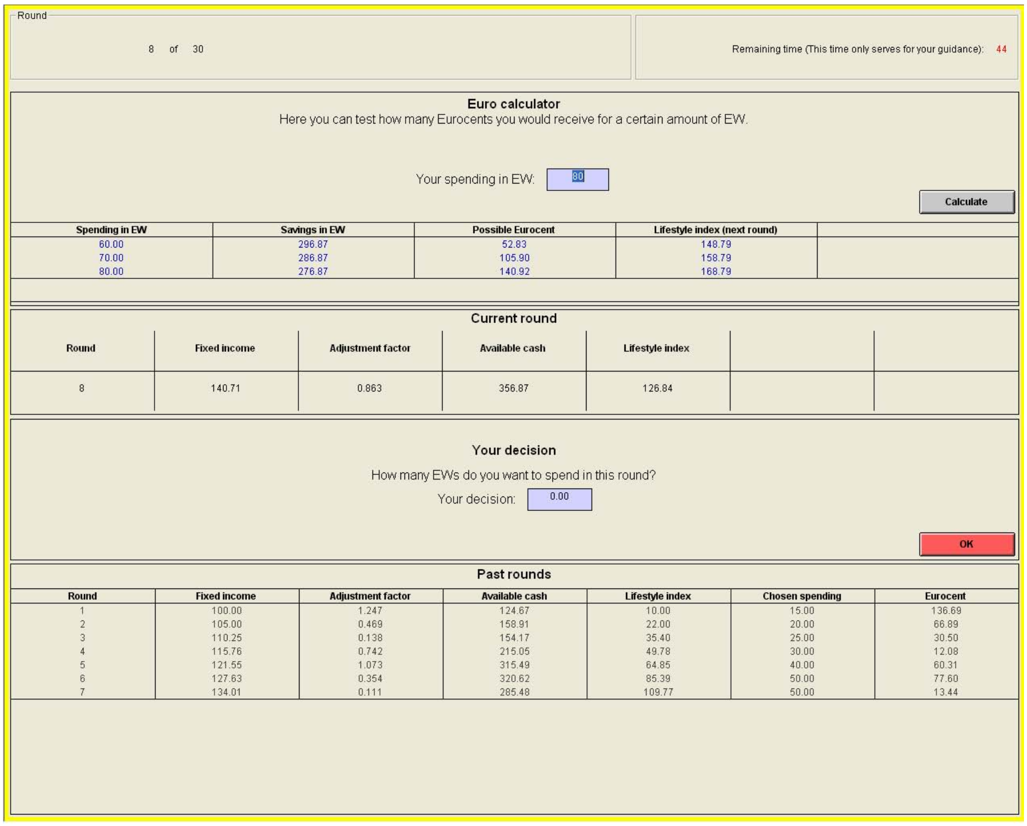


Figure 5.1: Screenshot of the decision screen

The second treatment dimension varies the framing of the random process generating subjects' income. The three framings identify changes in tax policy (TaxFrame), unemployment (JobFrame), or the general economic development (EconFrame) as the source of income shocks.

Finally, the communication of other subjects' beliefs about the income shock is used as a treatment variable. In treatment Belief, subjects are randomly assigned to groups each consisting of five subjects; this assignment is kept throughout the experiment. Before each of the 30 periods, subjects are informed about the beliefs of their fellow four group members.

The experiment was conducted computer-based and took place at the "Essen laboratory for experimental economics" (elfe) at the University of Duisburg-Essen, Germany, in October and November 2011. Participants were recruited via the program ORSEE (Greiner, 2004) and the attached subject pool. To program the experiment, the software z-Tree (Fischbacher, 2007) was used. In total, 11 sessions with up to 24 subjects each were conducted, leading to a total of 206 participants[109]. All participants were students from the University of Duisburg-Essen. The experiment lasted about 150 minutes and average payoff including the show-up fee of EUR 5.00 was EUR 27.39 (minimum EUR 0.10, maximum EUR 40.12).[110] Subjects were paid out one after the other to preserve anonymity.

5.4 Results

5.4.1 General picture

To get a general overview how the ex ante optimal solution looks like, Figure 5.2 illustrates the optimal consumption, cash-on-hand, and savings that are predicted for 30 periods. The stylized description of this is to save a lot (proportionately) in the early periods to build up a buffer stock of cash-on-hand. This way, consumption can be smoothed and no immediate reaction to bad draws in the future is necessary. At the very end, all savings are of course consumed.

The subjects in our experiment came surprisingly close to this prediction. Figure 5.3 shows the average of the subjects' actual consumption paths, the ex-ante optimal consumption, and the average of the respective conditionally optimal consumptions. Again, some stylized observations can be made: In the very first periods, overconsumption can be seen. A special case of this are some periods where the optimal consumption would have

[109]Of those, 4 subjects experienced technical problems and therefore are excluded from the analysis.

[110]If decisions by the participant would have lead to negative payoffs which exceeded the show-up fee, the participant was warned by a pop-up screen and could change her decision.

Figure 5.2: Ex ante optimal consumption, cash-on-hand, and savings

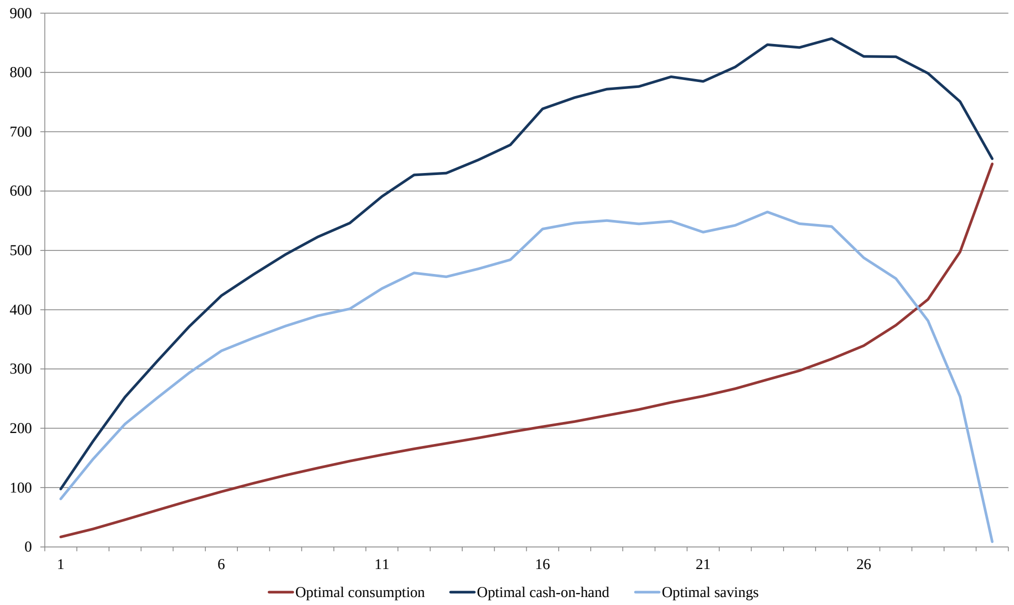
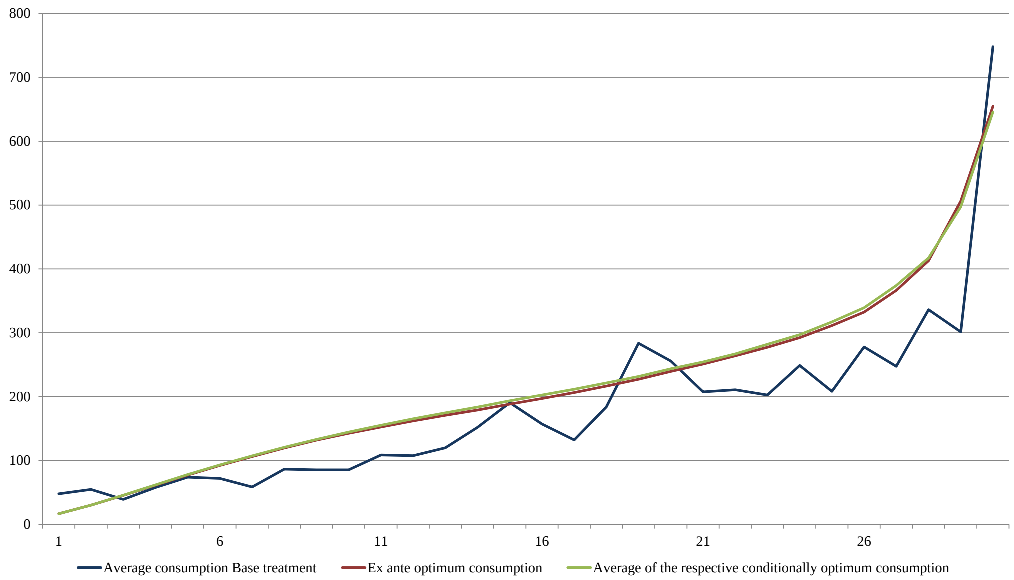


Figure 5.3: Average of actual consumption paths and conditional optima

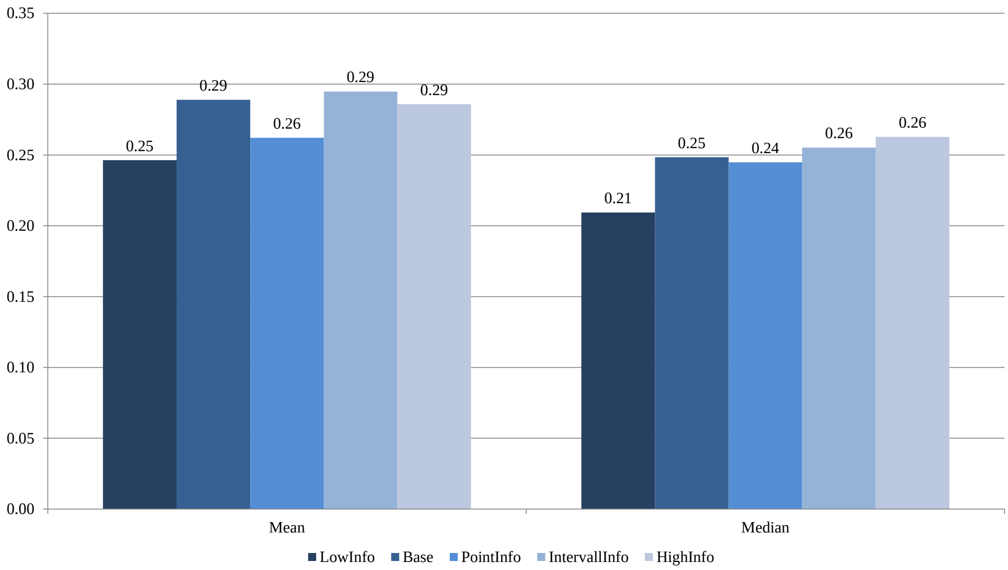


been 0 (recall that the negative utility one can incur is bounded from below) but the actual consumption is positive. In later rounds, the main difference is the overreaction to high income shocks. This way, the consumption smoothing is not as good as it would be in the conditional optimum. However, overall the actual consumption is surprisingly close to the actual chosen consumption. Still, the average deviation from the conditional optimum is significantly different from 0 in most rounds (only in rounds 14, 20, and 26 no significant difference can be found using a one sample t-test for each round).

5.4.2 The effect of information

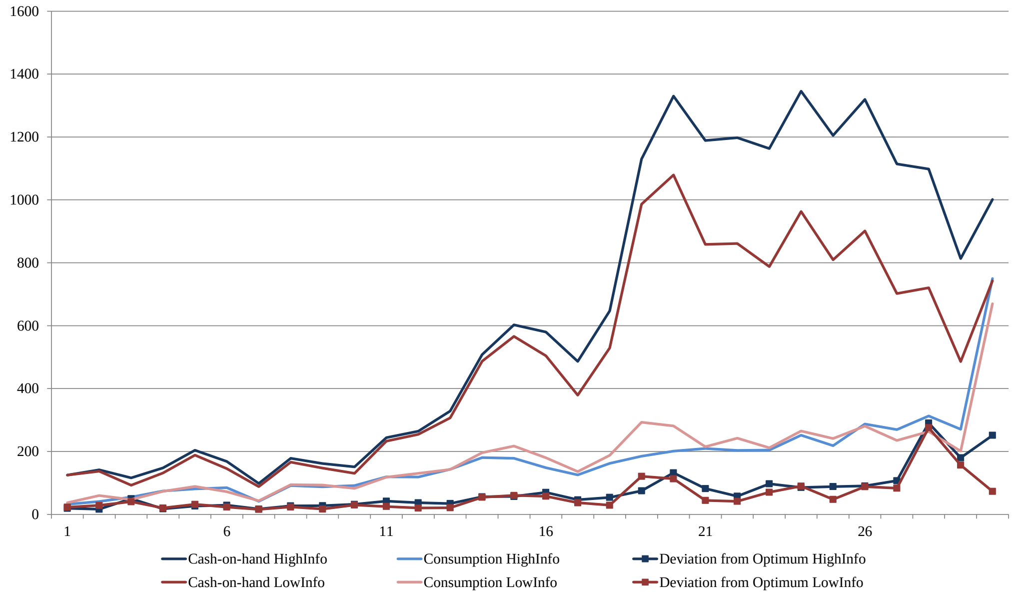
Given that subjects' consumption decisions are surprisingly close to the optimum, we test, whether different amounts of information about the stochastic income have an influence on this behavior. Figure 5.4 shows the average absolute deviation from the conditional optimum for each Info treatment and the Base treatment as a share of the fixed income.

Figure 5.4: Average absolute deviation from conditional optimum, Info and Base treatments



Using pairwise exact Mann-Whitney U tests to compare behavior between treatments reveals significant differences only between the LowInfo treatment and the HighInfo treatment ($p = 0.045$). But, instead of improving consumption decisions, more information leads to larger deviations from the conditional optimum, if at all. This observation is not consistent over all rounds: The treatment effect is (at least weakly) significant in rounds 2, 11, 12, 13, 18, and 21 ($p < 0.094$). It seems that very detailed information about the random process make subjects somewhat overconfident in their ability to determine the optimal consumption amount and, therefore, they tend to consume too much and save too little respectively.[111] Accordingly, the problem of undersaving (see above) is more sincere with more instead of less information. Figure 5.5 illustrates the actual consumption decisions, the cash-on-hand and the deviation from the conditional optimum observed in the two treatments for all 30 periods.

Figure 5.5: Consumption, cash-on-hand, and deviation from optimum

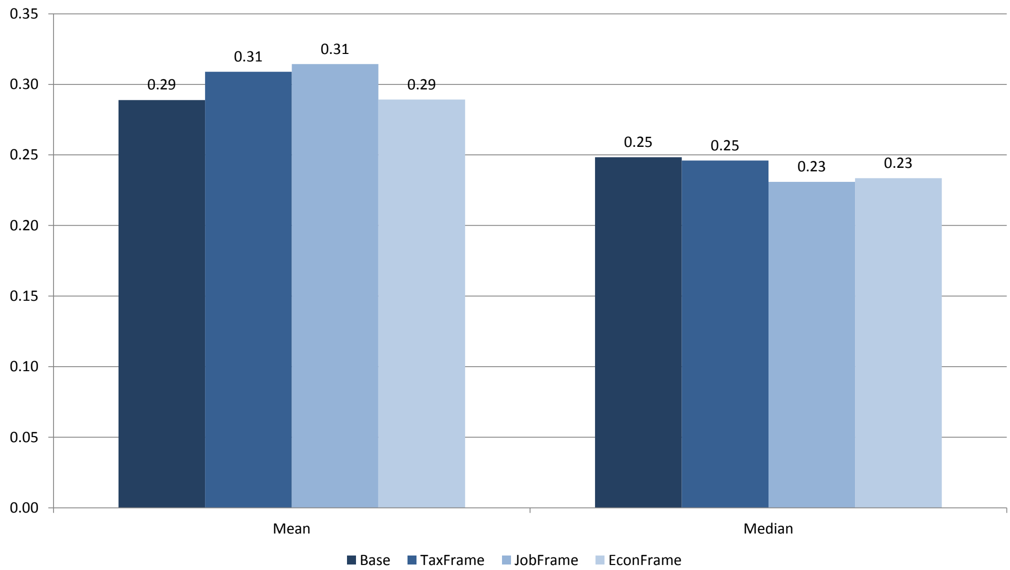


The previous results suggest that very detailed information rather produces a feeling of better understanding instead of being really understood by subjects. The question is then whether the framing of the random process and, thus, own experiences with income shocks triggered by this framing, do affect undersaving. We particularly focus on framings that (currently) have rather negative connotations and, accordingly, might induce subjects to consume more carefully and reduce undersaving. Figure 5.6 includes the average absolute deviation from the conditional optimum for each Frame treatment and the Base treatment as a fraction of fixed income. But neither the unemployment frame, nor the taxation frame, nor the general economic development frame of the random process significantly

[111] Overconfidence is a very robust finding (Among others, Kirchler and Maciejovsky 2002 provide an overview of several examples) and can broadly be classified into three different areas: miscalibration (an underestimation of the variance of random variables), unrealistically positive self-evaluations, and illusion of control (overestimation of personal success probability). For more details about these three areas see, e.g., Glaser et al. (2007, 2012). Menkoff et al. (2010) show that expert subjects - who usually have more information - exhibit a higher degree of overconfidence than laypeople, leading to worse decisions. Here, the higher amount of information available might also lead to more overconfidence and thus cause worse decisions.

affects behavior (p $>$ 0.841). Obviously the observed undersaving in our experiment is rather robust to experiences with income shocks made outside the laboratory.

Figure 5.6: Average absolute deviation from conditional optimum, framing treatments

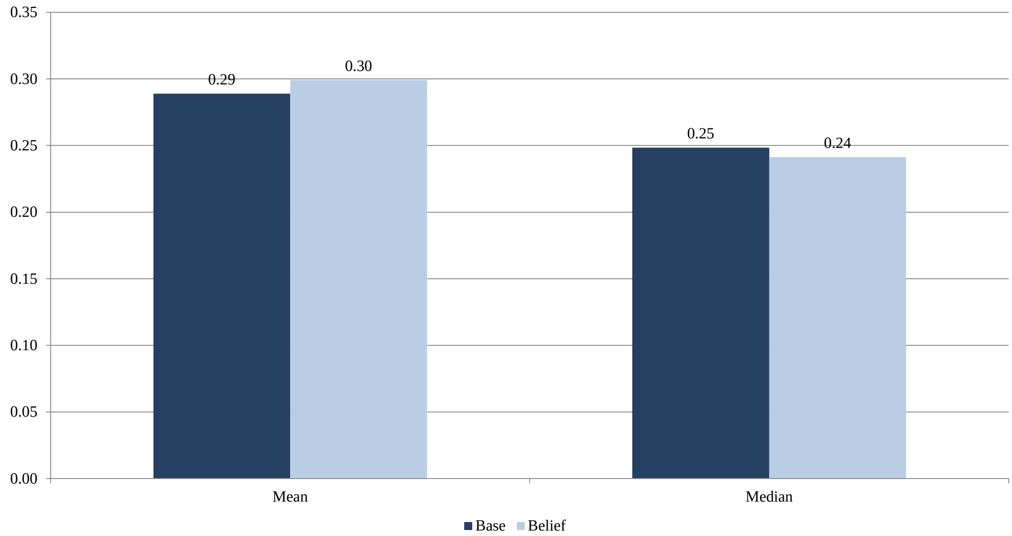


Brown et al. (2009) reported that social learning in the form of seeing other subjects' decision improves consumption behavior. Our belief treatment tests whether social learning can be already induced by seeing other subjects' beliefs about the next income shock. Figure 5.7 illustrates the aggregate data observed in the Belief and in the Base treatment. However, no significant difference between the Belief and the Base treatment can be observed (p = 0.575). In contrast to seeing others' decisions, subjects do not benefit when they are given access to the beliefs about future shocks.

5.4.3 Panel estimation

In order to control for confounding influences, a Random Effects panel estimation is conducted, using the log of absolute deviation from the conditional optimum as the dependent variable. An additional clustering at the group level is used for the participants in the Belief treatment who were informed about their group members' beliefs. Table 5.1 shows three different model specifications which add further explaining variables. The first specification includes only the treatment dummies with the baseline treatment as the reference treatment. The second one adds demographic variables (sex, age, a dummy for being an economics student, number of semesters, high school GPA), the number of test entries, and the subject's own belief. In addition, risk attitude, patience, and impulse control are

Figure 5.7: Average absolute deviation from conditional optimum, belief treatment



included.[112] The third and final specification also adds the period and the squared period to look for changes during the course of the experiment.

In all specifications, the treatment dummies are not significant (the Base treatment is used as a baseline). However, in specifications 2 and 3 a subject's belief, and number of test entries and in the full specification also the period and period squared, have a significant influence on behavior. The development of deviations over the thirty periods mirrors the results by Brown et al. (2009) as the decisions by the subjects actually get worse instead of better. This is in parts driven by the fact that the absolute scale of deviations just gets bigger as participants have a higher fixed income and more cash-on-hand. However, using the deviation as a share of fixed income instead yields the same qualitative results: deviations get larger in later rounds.

Looking at a subject's belief about the random adjustment factor in this round shows that a significant, slightly negative relation between the belief and the deviation exists. The existence of a relationship is somewhat puzzling, as the true adjustment factor for each round is drawn and shown to the subjects before they make their decision. Possibly, a subject's belief is an indicator for the anticipated general future development of the adjustment factor and, as such, influences behavior. Adding the beliefs of the other group members as explaining variables shows that none of the other beliefs has a significant influence on a subjects' behavior. Thus, subjects seem to realize that all group members

[112] These variables are evaluated on a 11-point scale with questions taken from the German Socio-Economic Panel (GSOEP).

Table 5.1: Random effects estimation results

Dependent variable: **Log of deviation**	**Specification 1**	**Specification 2**	**Specification 3**
(Constant)	3.312***	3.636***	2.219***
	(0.000)	(0.000)	(0.000)
Treatment 1	-0.166	-0.124	-0.127
	(0.883)	(0.364)	(0.356)
Treatment 3	-0.140	-0.114	-0.122
	(0.313)	(0.399)	(0.366)
Treatment 4	-0.004	0.003	-0.000
	(0.972)	(0.982)	(0.998)
Treatment 5	0.033	0.064	0.063
	(0.797)	(0.539)	(0.597)
Treatment 6	0.000	-0.016	-0.026
	(1.000)	(0.912)	(0.862)
Treatment 7	0.070	0.104	0.103
	(0.601)	(0.431)	(0.438)
Treatment 8	-0.052	-0.028	-0.028
	(0.698)	(0.836)	(0.832)
Treatment 9	-0.039	-0.021	-0.027
	(0.718)	(839)	(0.801)
Own belief	–	-0.001**	-0.001**
		(0.032)	(0.011)
Male dummy	–	-0.052	-0.050
		(0.507)	(0.518)
Age	–	-0.011	-0.011
		(0.167)	(0.143)
Semester	–	0.005	0.005
		(0.585)	(0.593)
Economics student dummy	–	0.003	0.021
		(0.110)	(0.100)
High school GPA	–	0.066	0.060
		(0.261)	(0.271)
Patience	–	-0.011	-0.011
		(0.497)	(0.494)
Impulse control	–	-0.022	-0.022
		(0.146)	(0.138)
Risk attitude	–	0.030	0.030
		(0.201)	(0.208)
Number of test entries	–	-0.043***	-0.051***
		(0.007)	(0.000)
Period	–	–	0.216***
			(0.000)
Period squared	–		-0.006***
			(0.000)

Notes: Treatment 2 is the baseline for the treatment dummies. P-values in parentheses. Additional clustering at group level.
Significance levels: *10% **5% ***1%
Source: Own calculations.

have the same information about the random adjustment factor and there is no need to take others' beliefs into account for their own consumption decision.

The effect of the number of test entries goes in the expected direction: The more entries in the Euro calculator a subject makes, the smaller is her deviation from the conditional optimum. However, the effect is pretty small in economic terms. Of course, one has to be careful in interpreting this significance, as it does not indicate a causal effect from number of test entries to quality of the decision. Instead, participants who simply think about the problem for a longer time might use both more test entries and make better decisions.

5.4.4 Heuristics

In general, our results reveal that even with little information, consumption decisions are quite close to the optimum. Given that more information, if at all, rather increases deviations to the optimum level raises some doubts that subjects calculate the optimal solution. Instead, it might be that they rather use a naïve heuristic basing their consumption decision on easier accessible variables like cash-on-hand or fixed income. To identify such simple decision rules, we run a regression (pooling all subjects) which explains the actual consumption decision by the predicted consumption, the available cash-on-hand, and the fixed income. Table 5.2 shows the results from this regression.

Table 5.2: OLS estimation results for actual consumption

Dependent variable: Consumption	
(Constant)	63.868
	(0.335)
Conditional Optimum	0.273*
	(0.059)
Cash-on-hand	0.045
	(0.247)
Fixed Income	-0.419
	(0.460)

Notes: P-values in parentheses.
Significance levels: *10% **5% ***1%
Source: Own calculations.

The regression results indicate that subjects do not consistently align their consumption to their cash-on-hand or their fixed income. Instead, the optimal consumption is weakly significantly correlated to their consumption decision. This observation suggests that subjects at least try to calculate the optimal decision though they are not able to account for all information necessary to determine this optimum.

5.5 Conclusion and discussion

The paper at hand uses a laboratory experiment to analyze the effect of different information conditions on consumption and savings behavior. The main results are twofold: First, the consumptions are - given the complex environment - pretty close to optimal decisions. Second, differing information conditions hardly affect the quality of the decisions. Especially the second result is somewhat surprising: The Info treatments tested here vary substantially from giving subjects almost no information at all about the random element of their future income to giving a full specification of the random term including example draws and the density function. Instead, if at all, more information seems to increase the deviation from the optimum. This suggests that subjects do not really use this information in the way assumed by standard models (i.e. thinking about their expected future utility) but instead employ some other way of dealing with the difficult decision situation. However, we can exclude simple decision rules that base the decision on easier accessible variables as a possible explanation. The observation that a considerable number of subjects use the Euro calculator (87% percent of the subjects use it in more than 25 periods) rather suggests that subjects at least try to calculate the optimum. Moreover, the fact that testing different consumption levels has a positive influence on decision quality implies that subjects may benefit even more from "learning by doing" than from thinking the model through right from the start. This mirrors the results by Brown et al. (2009), who repeat the "lifetime" of a subject seven times and find that decisions continually improve. It is also in line with our finding, that subjects do not benefit from just seeing the others' beliefs about income shocks - which still requires some calculation to determine own consumption, but rather need to see the decisions made by others to improve consumption behavior (as reported by Brown et al., 2009). Still, the exact method people use to arrive at their actual spending decisions is not clear. Future research might try to open the "black box" of the consumption decision and analyze how consumers are able to consistently make good decisions.

Conclusion

Communication and group behavior are ubiquitous aspects of decision making in economics. The thesis at hand presents five studies which investigate the role of these two aspects from different angles. While the studies differ in many dimensions, all point to the importance of the elements of group behavior and especially communication. This is both true for communication as a result itself and as a factor influencing results. While many economic theories and models disregard potential differences between individual and groups and treat communication only in a very stylized way, the examples presented here highlight situations where a closer look at actual communication and group dynamics are worthwhile to understand the observed behavior.

However, it is not feasible to point to one deciding aspect or factor of communication in a group and its influence on behavior. Instead, it is of critical importance that the context and the type of communication is taken into account as well. This is also true when designing institutions which either aim to influence communication or use communication to influence results. Chapters 3 shows that a conscious effort to influence communication - in this case taken as the transmission of information from one person to another person - is not always easy. Similarly, 4 shows that exogenously choosing communication channels to make the exchange of information faster is not trivial. Instead, one might have to deal with unintended consequences, both in the case of monetary incentives for increase information transmission and in the case of allowing only certain communication venues. The question of possible unintended consequences also plays a role in chapter 5, where more information does not as expected improve decision making. Again, the context of the decision has to be taken into account when assessing possible outcomes.

Apart from such a content-driven view, using experimental methods to look at communication content in small groups is also interesting from a methodological standpoint. When directly analyzing the communication content, the “black box” of decision making can be opened to some extent and the motives behind observed decisions might be deduced. This is especially interesting in situations where several competing motives behind one observed decision are possible. Examples for this can be found in chapters 1 and 3, where the communication content (from the chat or the video, respectively) is recorded and analyzed. In combination with the analysis of the final decision, the communication content can yield valuable insights into the individuals' decision process.

Stepping back from a purely academic point of view, it is intuitively clear that communication and group behavior plays an important role in everyday decision making. This importance ranges from the very small decision being made every day to major life decisions such as job search or long term finance planning. Following this, a deeper understanding of communication and group behavior in economic decision making is a worthwhile aim.

List of figures

List of tables

Bibliography

Adams, S. J., J. S. Heywood, and R. Rothstein (2009): *Teachers, Performance Pay, and Accountability: What Education Should Learn from Other Sectors*, Economic Policy Institute.

Akerlof, G. A. (1982): Labor contracts as partial gift exchange, *The Quarterly Journal of Economics*, 97, 543–569.

Akerlof, G. A. and J. L. Yellen (1990): The fair Wage-Effort hypothesis and unemployment, *The Quarterly Journal of Economics*, 105, 255.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutstroem (2008): Risk aversion in game shows, in *Risk Aversion in Experiments. Research in Experimental Economics*, ed. by G. W. H. James C. Cox, Bingley, UK: Emerald Publ., 359–404.

Andreoni, J. and J. H. Miller (2002): Giving according to GARP: an experimental test of the consistency of preferences for altruism, *Econometrica*, 70, 737–753.

Asheim, B. T. and L. Coenen (2005): Knowledge bases and regional innovation systems: Comparing nordic clusters, *Research Policy*, 34, 1173–1190.

Attanasio, O. P. and G. Weber (1989): Intertemporal substitution, risk aversion and the euler equation for consumption, *The Economic Journal*, 99, 59–73.

Baker II, R. J., S. K. Laury, and A. W. Williams (2008): Comparing small-group and individual behavior in lottery-choice experiments, *Southern Economic Journal*, 75, 367–382.

Bala, V. and S. Goyal (2000): A noncooperative model of network formation, *Econometrica*, 68, 1181–1229.

Baltussen, G., T. Post, M. J. V. D. Assem, and P. P. Wakker (2011): Random incentive systems in a dynamic choice experiment, *SSRN eLibrary*.

Bardsley, N. (2008): Dictator game giving: altruism or artefact?, *Experimental Economics*, 11, 122–133.

Barsky, R. B. (1997): Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study, *The Quarterly Journal of Economics*, 112, 537–79.

Bavelas, A. (1950): Communication patterns in task-oriented groups, *Journal of the Acoustical Society of America*, 22, 725–730.

Beetsma, R. M. W. J. and P. C. Schotman (2001): Measuring risk attitudes in a natural experiment: Data from the television game show lingo, *The Economic Journal*, 111, 821–848.

Bernhard, H., E. Fehr, and U. Fischbacher (2006): Group affiliation and altruistic norm enforcement, *The American Economic Review*, 96, 217–221.

Blinder, A. S. (2007): Monetary policy by committee: Why and how?, *European Journal of Political Economy*, 23, 106–123.

Bliss, R. T., M. E. Potter, and C. Schwarz (2011): Decision making and risk aversion in the cash cab, Working paper series, SSRN eLibrary.

Bloch, F. and M. O. Jackson (2007): The formation of networks with transfers among players, *Journal of Economic Theory*, 133, 83–110.

Bolton, G. E. and A. Ockenfels (2000): ERC: a theory of equity, reciprocity, and competition, *The American Economic Review*, 90, 166–193.

Bornstein, G., T. Kugler, and A. Ziegelmeyer (2004): Individual and group decisions in the centipede game: Are groups more "rational" players?, *Journal of Experimental Social Psychology*, 40, 599–605.

Bornstein, G. and I. Yaniv (1998): Individual and group behavior in the ultimatum game: Are groups more “rational” players?, *Experimental Economics*, 1, 101–108.

Boun My, K., M. Willinger, and A. Ziegelmeyer (1999): Global versus local interaction in coordination games: an experimental investigation, Tech. rep., Bureau d'Economie Theorique et Applique, UDS, Strasbourg.

Brooks, R., R. Faff, D. Mulino, and R. Scheelings (2009): Deal or no deal, that is the question: The impact of increasing stakes and framing effects on Decision-Making under risk, *International Review of Finance*, 9, 27–50.

Brosig, J., T. Heinrich, T. Riechmann, R. Schoeb, and J. Weimann (2010): Laying off or not? the influence of framing and economics education, *International Review of Economic Education*, 9, 44–55.

Brosig, J., A. Ockenfels, and J. Weimann (2003): The effect of communication media on cooperation, *German Economic Review*, 4, 217–241.

Brosig-Koch, J., C. Helbach, A. Ockenfels, and J. Weimann (2011): Still different after all these years: Solidarity behavior in east and west germany, *Journal of Public Economics*, 95, 1373–1376.

Brown, A. L., Z. E. Chua, and C. F. Camerer (2009): Learning and visceral temptation in dynamic saving experiments, *Quarterly Journal of Economics*, 124, 197–231.

Browning, M. and A. Lusardi (1996): Household saving: Micro theories and micro facts, *Journal of Economic Literature*, 34, 1797–1855.

Bull, C., A. Schotter, and K. Weigelt (1987): Tournaments and piece rates: An experimental study, *Journal of Political Economy*, 95, 1–33.

Camerer, C. F. and R. M. Hogarth (1999): The effects of financial incentives in experiments: A review and Capital-Labor-Production framework, *Journal of Risk and Uncertainty*, 19, 7–42.

Carbone, E. and J. D. Hey (2004): The effect of unemployment on consumption: an experimental analysis, *Economic Journal*, 114, 660–683.

Carpenter, J. P., S. Kariv, and A. Schotter (2010): Network architecture and mutual monitoring in public goods experiments, *SSRN eLibrary*.

Carroll, C. (2001): A theory of the consumption function, with and without liquidity constraints, *Journal of Economic Perspectives*, 15, 23–45.

Carroll, C. D. (1997): Buffer-stock saving and the life cycle/permanent income hypothesis, *The Quarterly Journal of Economics*, 112, 1–55.

Carroll, C. D., J. Overland, and D. N. Weil (2000): Saving and growth with habit formation, *The American Economic Review*, 90, 341–355.

Cason, T. N. and V. Mui (1997): A laboratory study in group polarisation in the team dictator game, *The Economic Journal*, 107, 1465–1483.

Charness, G., D. Levin, and E. Karni (2007a): Individual and group decision making under risk: An experimental study of bayesian updating and violations of first-order stochastic dominance, *Journal of Risk and Uncertainty*, 35, 129–148.

Charness, G., L. Rigotti, and A. Rustichini (2007b): Individual behavior and group membership, *American Economic Review*, 97, 1340–1352.

Chaudhuri, A., S. Graziano, and P. Maitra (2006): Social learning and norms in a public goods experiment with Inter-Generational advice, *Review of Economic Studies*, 73, 357–380.

Chen, Y. and S. X. Li (2009): Group identity and social preferences, *American Economic Review*, 99, 431–457.

Cohen, A. and L. Einav (2007): Estimating risk preferences from deductible choice, *American Economic Review*, 97, 745–788.

Cooper, D. J. and J. H. Kagel (2005): Are two heads better than one? team versus individual play in signaling games, *The American Economic Review*, 95, 477–509.

Cox, J. C. (2002): Trust, reciprocity and Other-Regarding preferences: Groups vs. individuals and males vs. females, in *Experimental Business Research*, ed. by R. Zwick and A. Rapoport, Springer, vol. 1, 331–350.

Crawford, V. (1998): A survey of experiments on communication via cheap talk, *Journal of Economic Theory*, 78, 286–298.

Croson, R. and S. Gächter (2010): The science of experimental economics, *Journal of Economic Behavior & Organization*, 73, 122–131.

De Roos, N. and Y. Sarafidis (2009): Desision making under risk in deal or no deal, *Journal of Applied Econometrics*.

Deaton, A. (1992): *Understanding Consumption*, Oxford University Press.

Deaton, A. S. (1991): Saving and liquidity constraints, *Econometrica*, 59, 1221–48.

Dohmen, T. and A. Falk (2010): You get what you pay for: Incentives and selection in the education system, *The Economic Journal*, 120, F256–F271.

Dohmen, T., A. Falk, D. Huffman, F. Marklein, and U. Sunde (2009): Biased probability judgment: Evidence of incidence and relationship to economic outcomes from a representative sample, *Journal of Economic Behavior & Organization*, 72, 903–915.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): Individual risk attitudes: Measurement, determinants, and behavioral consequences, *Journal of the European Economic Association*, 9, 522–550.

Donkers, B., B. Melenberg, and A. Van Soest (2001): Estimating risk attitudes using lotteries: A large sample approach, *Journal of Risk and Uncertainty*, 22, 165–195.

Eberts, R., K. Hollenbeck, and J. Stone (2002): Teacher performance incentives and student outcomes, *The Journal of Human Resources*, 37, 913–927.

Eckel, C. C. and P. J. Grossman (2005): Managing diversity by creating team identity, *Journal of Economic Behavior & Organization*, 58, 371 – 392.

Edquist, C. (2005): Systems of innovation: Perspectives and challenges, in *The Oxford Handbook of Innovation*, ed. by J. Fagerberg, D. C. Mowery, and R. R. Nelson, Oxford University Press, vol. 1, 181–208.

Falk, A. and J. J. Heckman (2009): Lab experiments are a major source of knowledge in the social sciences, *Science*, 326, 535–538.

Falk, A. and M. Kosfeld (2003): It's all about connections: Evidence on network formation, Tech. rep., C.E.P.R. Discussion Papers.

Fehr, E., U. Fischbacher, B. V. Rosenbladt, J. Schupp, and G. G. Wagner (2002): A Nation-Wide laboratory - examining trust and trustworthiness by integrating behavioral experiments into representative surveys, *Schmollers Jahrbuch*, 122, 519–542.

Fehr, E., E. Kirchler, A. Weichbold, and S. Gächter (1998): When social norms overpower competition: Gift exchange in experimental labor markets, *Journal of Labor Economics*, 16, 324–351.

FEHR, E. AND K. M. SCHMIDT (1999): A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics*, 114, 817–868.

FEHR-DUDA, H., A. BRUHIN, T. EPPER, AND R. SCHUBERT (2010): Rationality on the rise: Why relative risk aversion increases with stake size, *Journal of Risk and Uncertainty*, 40, 147–180.

FEHR-DUDA, H., M. DE GENNARO, AND R. SCHUBERT (2006): Gender, financial risk, and probability weights, *Theory and Decision*, 60, 283–313.

FEHR-DUDA, H., T. EPPER, A. BRUHIN, AND R. SCHUBERT (2011): Risk and rationality: The effects of mood and decision rules on probability weighting, *Journal of Economic Behavior & Organization*, 78, 14 – 24.

FIGLIO, D. N. AND L. W. KENNY (2007): Individual teacher incentives and student performance, *Journal of Public Economics*, 91, 901–914.

FISCHBACHER, U. (2007): z-Tree: zurich toolbox for ready-made economic experiments, *Experimental Economics*, 10, 171–178.

FISCHBACHER, U. AND F. HEUSI (2008): Lies in disguise. an experimental study on cheating, Tech. rep., Thurgauer Wirtschaftsinstitut, Universität Konstanz.

FLOOD, M. M. (1958): Some experimental games, *Management Science*, 5, 5–26.

FREY, B. S. (1997): *Markt und Motivation. Wie ökonomische Anreize die (Arbeits-) Moral verdrängen*, München: Vahlen.

FREY, B. S. AND R. JEGEN (2001): Motivation crowding theory, *Journal of Economic Surveys*, 15, 589–611.

FULLENKAMP, C., R. TENORIO, AND R. BATTALIO (2003): Assessing individual risk attitudes using field data from lottery games, *The Review of Economics and Statistics*, 85, 218–226.

GÄCHTER, S., G. VON KROGH, AND S. HAEFLIGER (2010): Initiating private-collective innovation: The fragility of knowledge sharing, *Research Policy*, 39, 893–906.

GEE, C. (2007): Risky choice and Type-Uncertainty in “Deal or no deal?”, Cambridge Working Papers in Economics 0758, Faculty of Economics, University of Cambridge.

GEROSKI, P. A. (2000): Models of technology diffusion, *Research Policy*, 29, 603–625.

GERTNER, R. (1993): Game shows and economic behavior: Risk-Taking on "Card sharks", *The Quarterly Journal of Economics*, 108, 507–521.

GIEBE, T., T. GREBE, AND E. WOLFSTETTER (2006): How to allocate R&D (and other) subsidies: An experimentally tested policy recommendation, *Research Policy*, 35, 1261–1272.

GLASER, M., T. LANGER, AND M. WEBER (2007): On the trend recognition and forecasting ability of professional traders, *Decision Analysis*, 4, 176–193.

GLASER, M., M. WEBER, AND T. LANGER (2012): True overconfidence in interval estimates: Evidence based on a new measure of miscalibration, *Journal of Behavioral Decision Making*, forthcoming.

GLEWWE, P., N. ILIAS, AND M. KREMER (2010): Teacher incentives, *American Economic Journal: Applied Economics*, 2, 205–227.

GNEEZY, U. AND A. RUSTICHINI (2000): Pay enough or don't pay at all, *Quarterly Journal of Economics*, 115, 791–810.

GOEREE, J. K., A. RIEDL, AND A. ULE (2009): In search of stars: Network formation among heterogeneous agents, *Games and Economic Behavior*, 67, 445–466.

GOETTE, L., D. HUFFMAN, AND S. MEIER (2006): The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups, *American Economic Review*, 96, 212–216.

GRANOVETTER, M. S. (1973): The strength of weak ties, *American Journal of Sociology*, 78, 1360–1380.

GREINER, B. (2004): An online recruitment system for economic experiments, in *Forschung und Wissenschaftliches Rechnen 2003. GWDG Bericht 63*, ed. by K. Kremer and V. Macho, 79–93.

GUETZKOW, H. AND H. A. SIMON (1955): The impact of certain communication nets upon organization and performance in Task-Oriented groups, *Management Science*, 1, 233–250.

HALL, R. E. (1988): Intertemporal substitution in consumption, *Journal of Political Economy*, 96, 339–57.

Harbring, C. and B. Irlenbusch (2003): An experimental study on tournament design, *Labour Economics*, 10, 443–464.

Hargreaves Heap, S. P. and A. Parikh (2005): The diffusion of ideas in the academy: A quantitative illustration from economics, *Research Policy*, 34, 1619–1632.

Harrison, G. W. and J. A. List (2004): Field experiments, *Journal of Economic Literature*, 42, 1009–1055.

Hartley, R., G. Lanot, and I. Walker (2006): Who really wants to be a millionaire? Estimates of risk aversion from gameshow data, Tech. rep., University of Warwick, Department of Economics.

Hersch, P. L. and G. S. McDougall (1997): Decision making under uncertainty when the stakes are high: Evidence from a lottery game show, *Southern Economic Journal*, 64, 75–84.

Hey, J. D. and V. Dardanoni (1987): Optimal consumption under uncertainty: An experimental investigation, *Economic Journal*, 98, 105–16.

Holt, C. A. and S. K. Laury (2002): Risk aversion and incentive effects, *American Economic Review*, 92, 1644–1655.

Jackson, M. O. and A. Wolinsky (1996): A strategic model of social and economic networks, *Journal of Economic Theory*, 71, 44–74.

Jianakoplos, N. A. and A. Bernasek (1998): Are women more risk averse?, *Economic Inquiry*, 36, 620–630.

Kahneman, D. (2003): Maps of bounded rationality: Psychology for behavioral economics, *American Economic Review*, 93, 1449–1475.

Kahneman, D. and A. Tversky (1979): Prospect theory: An analysis of decision under risk, *Econometrica*, 47, 263–291.

Keser, C., K. Ehrhart, and S. K. Berninghaus (1998): Coordination and local interaction: experimental evidence, *Economics Letters*, 58, 269–275.

Kessler, I. and J. Purcell (1992): Performance related pay: Objectives and application, *Human Resource Management Journal*, 2, 16–33.

KIRCHKAMP, O. AND R. NAGEL (2000): Local and group interaction in prisoners' dilemmas, Tech. rep., Sonderforschungsbereich 504, Universität Mannheim & Sonderforschungsbereich 504, University of Mannheim.

KIRCHLER, E. AND B. MACIEJOVSKY (2002): Simultaneous over- and underconfidence: Evidence from experimental asset markets, *Journal of Risk and Uncertainty*, 25, 65–85.

KNIGHT, F. H. (1921): *Risk, Uncertainty, and Profit*, Library of Economics and Liberty.

KOCHER, M. G. AND M. SUTTER (2008): The decision maker matters: Individual versus group behaviour in experimental Beauty-Contest games, *SSRN eLibrary*.

KOSFELD, M. (2004): Economic networks in the laboratory: A survey, *Review of Network Economics*, 3.

KUGLER, T., G. BORNSTEIN, M. G. KOCHER, AND M. SUTTER (2007): Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy, *Journal of Economic Psychology*, 28, 646–657.

LAVY, V. (2002): Evaluating the effect of teachers' group performance incentives on pupil achievement, *Journal of Political Economy*, 110, 1286–1317.

——— (2009): Performance pay and teachers' effort, productivity, and grading ethics, *The American Economic Review*, 99, 1979–2021.

LAZEAR, E. P. (2000): Performance pay and productivity, *The American Economic Review*, 90, 1346–1361.

LEAVITT, H. J. (1951): Some effects of certain communication patterns on group performance, *The Journal of Abnormal and Social Psychology*, 46, 38–50.

LEVITT, S. AND J. A. LIST (2006): What do laboratory experiments tell us about the real world?, *Journal of Economic Perspectives*, 21, 153–174.

LEVITT, S. D. AND J. A. LIST (2007): Viewpoint: On the generalizability of lab behaviour to the field, *Canadian Journal of Economics*, 40, 347–370.

LIST, J. (2006): Friend or foe? a natural experiment of the prisoner's dilemma, *The Review of Economics and Statistics*, 88, 463–471.

LIST, J. A. (2007): On the interpretation of giving in dictator games, *Journal of Political Economy*, 115, 482–493.

Luhan, W., M. Kocher, and M. Sutter (2009): Group polarization in the team dictator game reconsidered, *Experimental Economics*, 12, 26–41.

Martin, P., T. Mayer, and F. Mayneris (2011): Spatial concentration and plant-level productivity in france, *Journal of Urban Economics*, 69, 182–195.

Martins, P. S. (2009): Individual teacher incentives, student achievement and grade inflation, Tech. rep., Queen Mary, University of London, School of Business and Management, Centre for Globalisation Research.

Matsen, E. and B. Strom (2010): Dominated choices in a simple game with large stakes, *Experimental Economics*, 13, 99–119.

Mazar, N. and D. Ariely (2006): Dishonesty in everyday life and its policy implications, *Journal of Public Policy & Marketing*, 25, 117–126.

Menkoff, L., M. Schmeling, and U. Schmidt (2010): Overconfidence, experience, and professionalism: An experimental study, Working paper.

Metrick, A. (1995): A natural experiment in "Jeopardy!", *The American Economic Review*, 85, 240–253.

Moscovici, S. and M. Zavalloni (1969): The group as polarizer of attitudes, *Journal of Personality and Social Psychology*, 12, 125–135.

Muralidharan, K. and V. Sundararaman (2011): Teacher performance pay: Experimental evidence from india, *The Journal of Political Economy*, 119, 39–77.

Ockenfels, A. and A. E. Roth (2006): Late and multiple bidding in second price internet auctions: Theory and evidence concerning different rules for ending an auction, *Games and Economic Behavior*, 55, 297–320.

Ockenfels, A. and J. Weimann (1999): Types and patterns: an experimental East-West-German comparison of cooperation and solidarity, *Journal of Public Economics*, 71, 275–287.

Ortmann, A. and R. Hertwig (2002): The costs of deception: Evidence from psychology, *Experimental Economics*, 5, 111–131.

PEMBERTON, M. B., C. A. INSKO, AND J. SCHOPLER (1996): Memory for and experience of differential competitive behavior of individuals and groups, *Journal of Personality and Social Psychology*, 71, 953–966.

POKORNY, K. (2008): Pay but do not pay too much: An experimental study on the impact of incentives, *Journal of Economic Behavior & Organization*, 66, 251 – 264.

POST, T., M. J. VAN DEN ASSEM, G. BALTUSSEN, AND R. H. THALER (2008): Deal or no deal? decision making under risk in a Large-Payoff game show, *American Economic Review*, 98, 38–71.

PRATHER, L. J. AND K. L. MIDDLETON (2002): Are n+1 heads better than one?: The case of mutual fund managers, *Journal of Economic Behavior & Organization*, 47, 103–120.

PRENDERGAST, C. (1999): The provision of incentives in firms, *Journal of Economic Literature*, 37, 7–63.

RABIN, M. (1993): Incorporating fairness into game theory and economics, *The American Economic Review*, 83, 1281–1302.

——— (2000): Risk aversion and expected-utility theory: A calibration theorem, *Econometrica*, 68, 1281–1292.

REBACK, R. (2008): Teaching to the rating: School accountability and the distribution of student achievement, *Journal of Public Economics*, 92, 1394–1415.

ROCKENBACH, B., A. SADRIEH, AND B. MATHAUSCHEK (2007): Teams take the better risks, *Journal of Economic Behavior & Organization*, 63, 412–422.

ROTH, A. E. (1993): The early history of experimental economics, *Journal of the History of Economic Thought*, 15, 184–209.

——— (1995): Introduction to experimental economics, in *The Handbook of Experimental Economics*, ed. by J. H. Kagel and A. E. Roth, Princeton University Press, 3–110.

——— (2002): The economist as engineer: Game theory, experimentation, and computation as tools for design economics, *Econometrica*, 70, 1341–1378.

ROTH, E. (2010): Is experimental economics living up to its promise?

Rothgang, M. and B. Lageman (2011): Innovationspolitischer Mehrwert durch Vernetzung?: Cluster- und Netzwerkförderung als Politikinstrument auf Bundes- und Länderebene, *Vierteljahrshefte zur Wirtschaftsforschung*, 80, 143–166.

Rothgang, M., M. Peistrup, and B. Lageman (2011): Industrial collective research networks in germany: Structure, firm involvement and use of results, *Industry and Innovation*, 18, 393–414.

Rubinstein, A. (2006): A sceptic's comment on the study of economics, *The Economic Journal*, 116, C1–C9.

RWI, Joanneum Research, ISG, and Universität Jena (2011): Begleitende Evaluierung des Förderinstruments "Spitzencluster-Wettbewerb" des BMBF. Die Spitzencluster - Organisation, Positionierung im Innovationsgeschehen und Netzwerkbildung., Tech. rep.

Samuelson, L. (2005): Economic theory and experimental economics, *Journal of Economic Literature*, 43, 65–107.

Schotter, A. and B. Sopher (2007): Advice and behavior in intergenerational ultimatum games: An experimental approach, *Games and Economic Behavior*, 58, 365–393.

Schram, A. (2005): Artificiality: The tension between internal and external validity in economic experiments, *Journal of Economic Methodology*, 12, 225–237.

Schunk, D. (2009): What determines household saving behavior? an examination of saving motives and saving decisions, *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)*, 229, 467–491.

Selten, R. (1967): Die Strategiemethode zur Erforschung eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments, in *Beiträge zur experimentellen Wirtschaftsforschung*, ed. by H. Sauermann, Mohr Siebeck, 136–168.

Shearer, B. (2004): Piece rates, fixed wages and incentives: Evidence from a field experiment, *Review of Economic Studies*, 71, 513–534.

Siegel, S. and L. E. Fouraker (1960): *Bargaining and group decision making: Experiments in bilateral monopoly*, New York, NY, US: McGraw-Hill.

Soete, L., B. Verspagen, and B. Ter Weel (2010): Systems of innovation, in *Handbook in Economics, Economics of Innovation*, Amsterdam: Elsevier, vol. 2.

Sörensen, F., J. Mattsson, and J. Sundbo (2010): Experimental methods in innovation research, *Research Policy*, 39, 313–322.

Sutter, M. (2009): Individual behavior and group membership: Comment, *American Economic Review*, 99, 2247–2257.

Sutter, M., M. G. Kocher, and S. Strauss (2007): Individuals and teams in UMTS-License auctions, *SSRN eLibrary*.

Tanaka, T., C. F. Camerer, and Q. Nguyen (2006): Poverty, politics, and preferences: Field experiments and survey data from vietnam, Tech. rep., UCLA Department of Economics.

Thurstone, L. L. (1931): The indifference function, *The Journal of Social Psychology*, 2, 139–167.

van Dijk, F., J. Sonnemans, and F. van Winden (2001): Incentive systems in a real effort experiment, *European Economic Review*, 45, 187–214.

Watzlawick, P., J. H. Beavin, and D. D. Jackson (1967): *Pragmatics of Human Communication*, New York: W. W. Norton.

Woessmann, L. (2011): Cross-country evidence on teacher performance pay, *Economics of Education Review*, 30, 404–418.

Zeldes, S. P. (1989): Optimal consumption with stochastic income: Deviations from certainty equivalence, *The Quarterly Journal of Economics*, 104, 275–298.

Appendices

A.1 Appendix to Chapter 1

General Instructions[113]

Welcome to the experiment and thank you for participating! Please read the instructions carefully. Please do not communicate with the other participants from now on. If you have a question, please raise your hand! We will come to your seat and answer your question. If you do not follow these rules, you unfortunately have to stop the experiment and will not receive any payment.

During the experiment, you can earn money depending on your choices and the choices of the other participants.

The experiment consists of four independent parts. There is no connection between the decisions in each of the parts. Before every part you will receive separate, detailed instructions. All instructions are identical for all participants. At the end of the experiment, one of the four parts will be randomly chosen. Your payoff from this experiment will be the amount of money you earned in this part. In all four parts your interaction partner is randomly determined anew.

Types of participants

At the beginning of the experiment, all participants are randomly assigned to a type (type A or type B). One half of the participants is assigned to type A, the other half is assigned to type B. The assignment to type A or type B stays the same during the whole experiment.

The first part of the experiment starts now.

[113] This is the English translation of the originally German instructions.

First Part

Every type A participant is randomly assigned to one type B participant. You will not learn about the identity of your assigned participant, neither during nor after the experiment. Thus, your decisions are completely anonymous.

Basic Endowment

Every participant receives a basic endowment of EUR 12.

Decision

Every type A participant has to make a decision: Type A can either leave the basic endowments unchanged, he can take away money from the basic endowment of his assigned type B participant, or he can give money from his own endowment to his assigned type B participant. A possible transfer is only allowed in increments on EUR 1. This results in the following possible distributions:

Payoff Type A (EUR)	0	1	2	3	4	5	6	7	8	9	10	11	**12**	13	14	15	16	17	18	19	20	21	22	23	24
Payoff Type B (EUR)	24	23	22	21	20	19	18	17	16	15	14	13	**12**	11	10	9	8	7	6	5	4	3	2	1	0

Information

Afterwards, every type B participant will be informed about the decision his assigned type B participant has made.

Payoffs

At the end of the experiment, one of the four parts will be randomly chosen. The decision in this part will then determine the payoffs of the participants.

Second Part

All type A participants are randomly assigned to groups. Three type A participants form one group. In each three-person group, the numbers 1, 2, and 3 are randomly assigned to the participants. Again, one type B participant is randomly assigned to every type A participant.

Basic Endowment
Every participant receives a basic endowment of EUR 12.

Decision
Every type A participant has to make a decision: Type A can either leave the basic endowments unchanged, he can take away money from the basic endowment of his assigned type B participant, or he can give money from his own endowment to his assigned type B participant. A possible transfer is only allowed in increments on EUR 1. This results in the following possible distributions:

Payoff Type A (EUR)	0	1	2	3	4	5	6	7	8	9	10	11	**12**	13	14	15	16	17	18	19	20	21	22	23	24
Payoff Type B (EUR)	24	23	22	21	20	19	18	17	16	15	14	13	**12**	11	10	9	8	7	6	5	4	3	2	1	0

One group member is randomly chosen. The decision of this group member is binding for all three group members.

Information
Afterwards, every type B participant will be informed about the decision his assigned type B participant has made, respectively about the decision the randomly chosen group member has chosen.

Payoffs
At the end of the experiment, one of the four parts will be randomly chosen. The decision in this part will then determine the payoffs of the participants.

Third Part

Again, all type A participants are randomly assigned to groups. Three type A participants form one group. In each three-person group, the numbers 1, 2, and 3 are randomly assigned to the participants. Again, one type B participant is randomly assigned to every type A participant.

Basic Endowment

Every participant receives a basic endowment of EUR 12.

Decision

The group members can talk with each other for 5 minutes via electronic chat. There is no contact with the assigned type B participants. During the electronic chat, the messages of participant number 1 are marked with ``1", the messages of participant 2 are marked with ``2", and the messages of participant 3 are marked with “3”. This designation is only meant to distinguish the different group members. During the electronic chat, you are not allowed to say your name, or give other hints to your identity. Anonymity among the group members stays intact. The content of the electronic chat is recorded for subsequent analysis.

After the electronic chat, the participant with the number 3 has to make a decision: Type A can either leave the basic endowments unchanged, he can take away money from the basic endowment of his assigned type B participant, or he can give money from his own endowment to his assigned type B participant. A possible transfer is only allowed in increments on EUR 1. This results in the following possible distributions:

Payoff Type A (EUR)	0	1	2	3	4	5	6	7	8	9	10	11	**12**	13	14	15	16	17	18	19	20	21	22	23	24
Payoff Type B (EUR)	24	23	22	21	20	19	18	17	16	15	14	13	**12**	11	10	9	8	7	6	5	4	3	2	1	0

The decision of the participant with the number 3 is binding for all group members. After participant number 3 has made the decision, participants number 1 and 2 are asked how they would have decided in his place without knowing his actual decision. The decision by participant 1 and 2 has no influence on the payoffs.

Information

Afterwards, every type B participant will be informed about the decision his assigned type B participant has made, respectively about the decision participant number 3 has made for the group members.

Payoffs

At the end of the experiment, one of the four parts will be randomly chosen. The decision in this part will then determine the payoffs of the participants.

Fourth Part

Every type A participant is randomly assigned to one type B participant. You will not learn about the identity of your assigned participant, neither during nor after the experiment. Thus, your decisions are completely anonymous.

Basic Endowment

Every participant receives a basic endowment of EUR 12.

Decision

Every type A participant has to make a decision: Type A can either leave the basic endowments unchanged, he can take away money from the basic endowment of his assigned type B participant, or he can give money from his own endowment to his assigned type B participant. A possible transfer is only allowed in increments on EUR 1. This results in the following possible distributions:

Payoff Type A (EUR)	0	1	2	3	4	5	6	7	8	9	10	11	**12**	13	14	15	16	17	18	19	20	21	22	23	24
Payoff Type B (EUR)	24	23	22	21	20	19	18	17	16	15	14	13	**12**	11	10	9	8	7	6	5	4	3	2	1	0

Information

Afterwards, every type B participant will be informed about the decision his assigned type B participant has made.

Payoffs

At the end of the experiment, one of the four parts will be randomly chosen. The decision in this part will then determine the payoffs of the participants.

A.2 Appendix to Chapter 2

Table A.2.1: Variable definitions

Variable name	Definition/examples
Stake (/100)	Acquired winnings at the end of the "Quiz Taxi" ride in EUR, divided by 100 (when losing the third life, or when reaching the master question)
Stake per contestant (/100)	Acquired winnings at the end of the "Quiz Taxi" ride in EUR per group member, divided by 100 (when losing the third life, or when reaching the master question)
Distance	Distance from start to final destination in kilometers
Number of questions	Number of questions played
Lives left	Lives left at the end of the "Quiz Taxi" ride (0 - 3)
Wild card left	Telephone or passersby wild card not used during "Quiz Taxi" ride (dummy variable)
Share of right answers (in %)	Correct answers divided by number of questions in percent
Share of known answers (in %)	Known answers (no use of wild cards, no obvious guesses) divided by number of questions in percent
Streak	Number of consecutive correct answers directly before master question
2 passengers	Two contestants (dummy variable)
3 passengers	Three contestants (dummy variable)
Females only	Only female contestants (dummy variable)
Males only	Only male contestants (dummy variable)
Females and males	Female and male contestants (dummy variable)
Young contestants (under 30)	All contestants younger than 30 (dummy variable)
Middle age contestants (30 - 50)	All contestants between 30 and 50 (dummy variable)
Old contestants (above 50)	All contestants older than 50 (dummy variable)
Contestants of different age	Contestants of different age (dummy variable)
Migration background	At least one groups member appears to have a migration background, based on language skills and appearance (dummy variable)
DVD episode	Run comes from DVD, not from internet portal "maxdome.de" (dummy variable)
Discussion time	Length of discussion before the final decision to play the master question in seconds
Majority pro master	More contestants initially want to play master question (dummy variable)
Majority contra master	More contestants initially do not want to play master question (dummy variable)
Number of topics discussed	Number of arguments exchanged
Number of topics pro	Number of arguments that imply that master question should be played
Number of topics contra	Number of arguments that imply that master question should not be played
Humility/modesty	"Modesty is a virtue!" "Greed is not good¡' (dummy variable)
A lot of money	"That's good money for us." "That would be a whole monthly salary." (w.r.t. stake only, dummy variable)
Nothing to lose	"Well, we entered with nothing, so we can't really lose anything!" (dummy variable)
Provocation/encouragement by host	"You cowards!" "You did so well so far." (dummy variable)
Easy to answer master question	"It's not that hard, anyway." "It's not more difficult than the others." (only if mentioned by contestants, dummy variable)
Difficult to answer master question	"We've been so lucky so far." "We were not that good so far." (only if mentioned by contestants, dummy variable)

Source: Own transcripts

Table A.2.2: Variable means by topics discussed

	Humble		A lot of money		Nothing to lose		Provocation		Easy		Difficult	
	No	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**
Stake (/100)	8.10	8.26	8.06	8.44	8.26	7.18	8.20	7.96	8.21	7.13	8.13	8.06
Stake per contestant (/100)	3.78	3.72	3.75	3.85	3.83	3.36	3.85	3.60	3.82	3.21	3.76	3.79
Share of correct answers (in %)	84.43	84.10	83.85	87.08	84.44	83.98	83.65	85.88	84.36	84.61	84.25	84.94
Share of known answers (in %)	65.18	66.57	65.05	67.06	65.01	67.85	65.07	66.03	65.85	60.14	65.54	64.70
Discussion time (in seconds)	26.77	30.14	25.78	34.75	25.39	39.74	20.61	40.81	25.09	51.42	24.17	40.85
3 passengers	0.18	0.24	0.17	0.25	0.19	0.16	0.16	0.23	0.18	0.25	0.19	0.15
Females only	0.25	0.19	0.25	0.17	0.24	0.21	0.29	0.15	0.24	0.25	0.24	0.26
Males only	0.31	0.33	0.30	0.42	0.29	0.47	0.26	0.44	0.31	0.33	0.33	0.26
Observations	125	21	122	24	127	19	98	48	134	12	119	27

Source: Own calculations

Table A.2.3: Robustness checks: different estimation models
(Average marginal effects from specified regressions)

Dependent variable:		**Probit selection**			
Play master question (1=yes, 0=no)	**Probit**	**1st stage**	**2nd stage**	**OLS**	**Logit**
Stake (/100)	-0.044***	0.109***	-0.045*	-0.044***	-0.043***
	(0.016)	(0.007)	(0.027)	(0.016)	(0.016)
Share of correct answers (in %)	0.013**	–	–	0.014*	0.012*
	(0.006)			(0.007)	(0.006)
Wild card left	0.075	0.019	0.083	0.083	0.079
	(0.093)	(0.062)	(0.098)	(0.107)	(0.095)
3 passengers	0.071	-0.121*	0.044	0.078	0.077
	(0.106)	(0.063)	(0.122)	(0.114)	(0.113)
Females only	-0.138	-0.061	-0.147	-0.138	-0.141
	(0.095)	(0.062)	(0.103)	(0.094)	(0.099)
Males only	-0.033	-0.023	-0.037	-0.037	-0.036
	(0.090)	(0.057)	(0.095)	(0.097)	(0.093)
DVD episode	0.270**	0.061	0.291**	0.301*	0.267**
	(0.129)	(0.086)	(0.145)	(0.153)	(0.133)
Distance	–	-0.085***	–	–	–
		(0.017)			
Pseudo/Adjusted R2	0.09			0.07	0.09
Model p-value	0.027	0.41		0.014	0.049
Observations	146	256		146	146

Notes: Table reports average marginal effects from specified regressions, robust standard errors in parentheses. Selection model: Heckman maximum-likelihood estimator; first stage instrument for reaching the master question: distance to final destination. Significance levels: *10% **5% ***1%
Source: Own calculations

Table A.2.4: Robustness checks: different sets of covariates
(Average marginal effects from probit regressions; all groups)

Dependent variable: **Play master question (1=yes, 0=no)**	**Additional Covariates**	**Stake per contestant**	**Log of stake**	**Interaction stake * % correct**	**Share of known answers**	**Streak of correct answers**	**Log of discussion time**
Stake (/100)	-0.036**	–	–	-0.041**	-0.035**	-0.031**	-0.037**
	(0.016)			(0.016)	(0.017)	(0.015)	(0.016)
Stake per contestant (/100)	–	-0.070**	–	–	–	–	–
		(0.034)					
Log of stake	–	–	-0.040**	–	–	–	–
			(0.018)				
Share of correct answers (in %)	0.010	0.010	0.011	0.013**	–	–	0.011*
	(0.007)	(0.007)	(0.007)	(0.007)			(0.007)
Share of known answers (in %)	–	–	–	–	0.005	–	–
					(0.004)		
Streak of correct answers	–	–	–	–	–	0.017	–
						(0.014)	
Wild card left	0.065	0.068	0.067	0.063	0.027	0.066	0.072
	(0.095)	(0.095)	(0.095)	(0.094)	(0.097)	(0.094)	(0.094)
3 passengers	0.101	-0.008	0.104	0.121	0.092	0.076	0.087
	(0.102)	(0.098)	(0.102)	(0.101)	(0.104)	(0.101)	(0.103)
Females only	-0.110	-0.112	-0.106	-0.103	-0.110	-0.118	-0.107
	(0.093)	(0.093)	(0.093)	(0.094)	(0.094)	(0.094)	(0.095)
Males only	-0.021	-0.023	-0.019	-0.014	-0.014	-0.010	-0.031
	(0.088)	(0.088)	(0.088)	(0.088)	(0.087)	(0.088)	(0.088)
DVD episode	0.170	0.177	0.167	0.144	0.184	0.185	0.190
	(0.135)	(0.134)	(0.134)	(0.134)	(0.136)	(0.136)	(0.131)
Discussion time (in seconds)	0.005***	0.005***	0.005***	0.005***	0.006***	0.005***	–
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
Log of discussion time	–	–	–	–	–	–	0.007**
							(0.003)
Young contestants (under 30)	-0.060	-0.064	-0.062	-0.059	-0.032	-0.050	-0.059
	(0.081)	(0.081)	(0.081)	(0.081)	(0.080)	(0.080)	(0.081)
Old contestants (above 50)	-0.043	-0.048	-0.061	-0.067	0.020	-0.040	-0.061
	(0.155)	(0.156)	(0.152)	(0.152)	(0.162)	(0.156)	(0.158)
Migration background	0.029	0.025	0.022	0.009	0.027	0.018	0.028
	(0.109)	(0.110)	(0.110)	(0.111)	(0.112)	(0.110)	(0.110)
Pseudo R2	0.13	0.13	0.13	0.14	0.13	0.13	0.12
Model p-value	0.042	0.046	0.045	0.025	0.043	0.033	0.038
Observations	146	146	146	146	146	146	146

Notes: Table reports average marginal effects from probit regressions, robust standard errors in parentheses. Interaction term included in regression of column 4, corresponding marginal effect not reported. Significance levels: *10% **5% ***1%
Source: Own calculations

Table A.2.5: Robustness checks: different sets of covariates
(Average marginal effects from probit regressions; undecided groups only)

Dependent variable: **Play master question (1=yes, 0=no)**	**Additional Covariates**	**Stake per contestant**	**Log of stake**	**Interaction stake * % correct**	**Share of known answers**	**Streak of correct answers**	**Log of discussion time**
Stake (/100)	-0.037	–	–	-0.033	-0.017	-0.011	-0.036
	(0.028)			(0.028)	(0.025)	(0.025)	(0.028)
Stake per contestant (/100)	–	-0.074	–	–	–	–	–
		(0.060)					
Log of stake	–	–	-0.028	–	–	–	–
			(0.031)				
Share of correct answers (in %)	0.030**	0.029**	0.028**	0.027**	–	–	0.031**
	(0.013)	(0.013)	(0.013)	(0.013)			(0.013)
Share of known answers (in %)	–	–	–	–	0.011**	–	–
					(0.005)		
Streak of correct answers	–	–	–	–	–	0.037*	–
						(0.020)	
Wild card left	-0.071	-0.068	-0.071	-0.065	-0.115	-0.062	-0.084
	(0.117)	(0.118)	(0.118)	(0.117)	(0.130)	(0.128)	(0.115)
3 passengers	0.112	0.009	0.103	0.098	0.082	0.047	0.099
	(0.120)	(0.128)	(0.123)	(0.122)	(0.130)	(0.124)	(0.119)
Females only	-0.302**	-0.305**	-0.304**	-0.310**	-0.254*	-0.286*	-0.299**
	(0.142)	(0.143)	(0.144)	(0.141)	(0.134)	(0.149)	(0.143)
Males only	-0.106	-0.110	-0.103	-0.126	-0.096	-0.059	-0.106
	(0.116)	(0.115)	(0.117)	(0.118)	(0.127)	(0.127)	(0.115)
DVD episode	-0.009	-0.001	-0.005	0.001	0.006	0.016	-0.003
	(0.169)	(0.168)	(0.168)	(0.170)	(0.175)	(0.165)	(0.165)
Discussion time (in seconds)	0.006**	0.006**	0.006**	0.006**	0.007**	0.007**	–
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	
Log of discussion time	–	–	–	–	–	–	0.010**
							(0.004)
Young contestants (under 30)	-0.121	-0.130	-0.128	-0.129	-0.139	-0.098	-0.120
	(0.101)	(0.101)	(0.103)	(0.101)	(0.113)	(0.113)	(0.101)
Old contestants (above 50)	0.096	0.090	0.052	0.200	0.022	-0.038	0.081
	(0.267)	(0.267)	(0.256)	(0.287)	(0.266)	(0.295)	(0.259)
Migration background	0.211	0.197	0.192	0.278**	0.206	0.125	0.229*
	(0.135)	(0.134)	(0.133)	(0.130)	(0.142)	(0.156)	(0.139)
Pseudo R2	0.23	0.23	0.22	0.24	0.17	0.17	0.22
Model p-value	0.128	0.126	0.124	0.078	0.161	0.301	0.119
Observations	65	65	65	65	65	65	65

Notes: Table reports average marginal effects from probit regressions, robust standard errors in parentheses. Interaction term included in regression of column 4, corresponding marginal effect not reported. Significance levels: *10% **5% ***1%
Source: Own calculations

A.3 Appendix to Chapter 3

Table A.3.1: Descriptive statistics, individual level

Variable	Minimum	Maximum	Average	Standard deviation
Share of correct answers	0.27	1.43	0.818	0.209
Participants in the pupil role				
Age	19	42	23.920	3.953
Experience with card games (binary)	0	1	0.340	0.475
Experience with the game used	0	0	0	0
Experience with the language used	0	0	0	0
Teaching experience (binary)	0	1	0.510	0.501
Female (binary)	0	1	0.495	0.501
Last math grade	6	15	11.505	1.831
Last German grade	6	15	11.272	2.257
Risk attitude	0	10	5.540	2.233
Study time (in semesters)	1	23	4.980	3.779
Trust score	1	4	2.634	0.589
Number of correct answers	6	26	16.490	3.497
Total observations	194			
Participants in the instructor role				
Age	19	40	23.880	3.416
Experience with card games (binary)	0	1	0.371	0.484
Experience with the game used	0	1	0.010	0.101
Experience with the language used	0	1	0.010	0.101
Teaching experience (binary)	0	1	0.701	0.459
Female (binary)	0	1	0.526	0.501
Last math grade	6	15	11.622	1.924
Last German grade	6	15	10.778	2.236
Risk attitude	0	10	5.340	2.088
Study time (in semesters)	1	19	5.580	3.647
Last slide seen	21	43	37.890	4.784
Trust score	1	4	2.624	0.634
Number of correct answers	12	29	20.090	3.519
Total observations	194			

Source: Own calculations.

Table A.3.2: Video coding variables

Variable name	**Description**
Time total	Time (in seconds) the instructor uses to teach
Time Pizza	Time (in seconds) the instructor uses to teach the topic "Pizza Bäcker"
Time Lojban	Time (in seconds) the instructor uses to teach the topic "Lojban"
Number switches	Number of times the instructor switches the topic
First topic	Which topic does the instructor start with (0="Pizzabäcker", 1="Lojban")
Material pupil	Does the instructor indicate that the pupil also has the material? (0=no, 1=yes, 2=instructor is not sure)
Motivation	Does the instructor motivate the pupil? (0=no, 1=yes)
Payoff	Does the instructor talk about the experiment's payoff rule? (0=no, 1=yes)
Own effort	Does the instructor comment on her own teaching performance? (0=no, 1=yes)
Time teaching	Does the instructor mention the time constraint during teaching? (0=no, 1=yes)
Time presentation	Does the instructor mention the time constraint during the presentation? (0=no, 1=yes)
Summary	Does the instructor provide a summary for one or both of the topics? (0=no, 1=yes)
Both topics	Does the instructor mention both topics at the beginning of the teaching period? (0=no, 1=yes)
Cut off	Is the instructor cut off in the middle of teaching or does she finish by herself? (0=no, 1=yes)
We	Does the instructor use the pronoun "we" (or "us")? (0=no, 1=yes)
Direct address	Does the instructor adress the pupil directly? (0=no, 1=yes)
Misunderstood	Does the instructor misunderstand the situation (e.g. expecting the pupil to talk to him)? (0=no, 1=yes)
Comment	Does the instructor comment the situation (e.g. saying that it is strange not to hear the pupil)? (0=no, 1=yes)
Camera	Does the camera point at the instructor's face? (0=no, 1=yes)
Eye contact	Does the instructor establish eye contact with the pupil? (0=no, 1=yes)
Speed	How fast is the instructor's rate of speech? (0=slow, 1=average, 2=high)
Material "Pizzabäcker"	Does the instructor hold material (playing cards) in front of the camera? (0=no, 1=yes)
Examples "Pizzabäcker"	How many examples does the instructor use to explain the game?
Mistakes "Pizzabäcker"	How many mistakes does the instructor make when explaining the game?
Material "Lojban"	Does the instructor hold metarial (example sheet) in front of the camera? (0=no, 1=yes)
Examples "Lojban"	How many examples does the instructor use to explain the language?
Mistakes "Lojban"	How many mistakes does the instructor make when explaining the language?
Question n	How many answering possibilities from question n can the pupil exclude from the instructor's explanations? (if the pupil is led to believe the wrong answer is correct, this is coded as 4.) $n \in [1;30]$.

Figure A.3.1: Average ratio of transferred knowledge of the topic "card game" by treatment

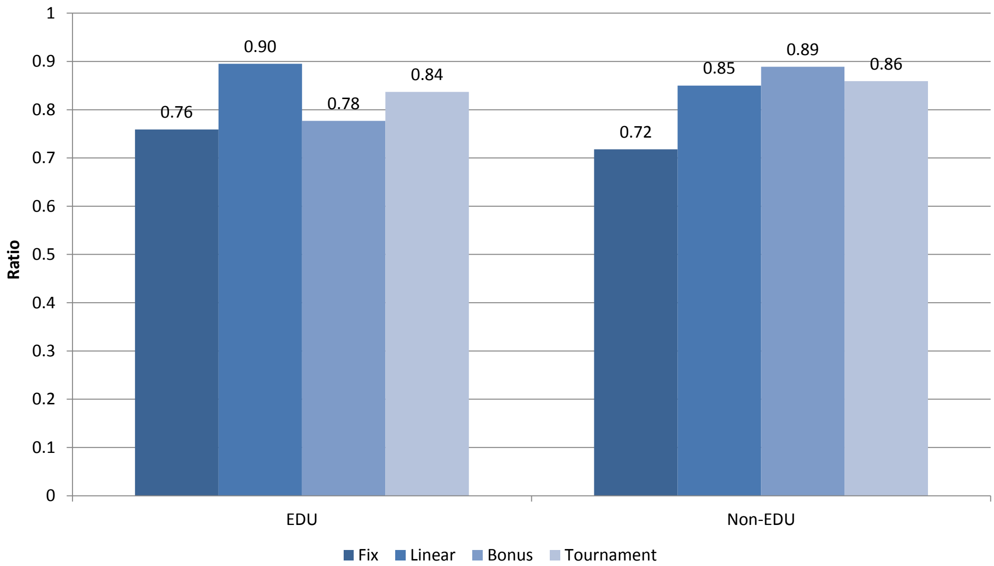
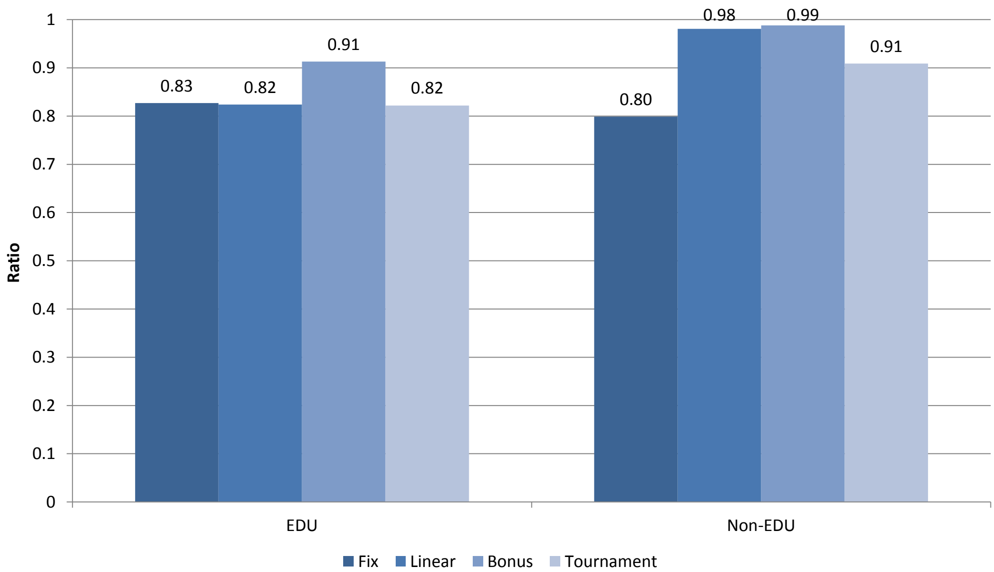


Figure A.3.2: Average ratio of transferred knowledge of the topic "artificial language" by treatment



Instructions[114]

Welcome to the experiment!

You are participating at a study of decision making behavior in the context of experimental economics. During the study you and the other participants will be asked to make decisions. You can earn money with this study. How much money you earn is dependent on the course of the experiment. You will receive detailed instructions about this in the following. All participants are paid in cash directly after the experiment one by one. To assure this, please remain seated after the experiment until your cabin number is called.

During the course of the experiment, no participant will receive information about the other participants' identity. All decisions are therefore made anonymously.

Should you have any questions before the start of the experiment, please ask an employee of the laboratory. He will come to your place and help you. **Any communication with the other participants during the experiment is only allowed when explicitly prompted; breaking this rule will lead to an immediate exclusion from the experiment.**

[114]These instructions are translated from the original, German instructions. Additional materials are available from the authors upon request.

Instructions

The experiment consists of three parts. In the following, you will receive detailed information about these. Please read the instructions carefully and thoroughly and click the start button on the screen **only after you have clarified all possible questions**. **After that, further questions cannot be answered any more.** In this experiment, you are either a **"teacher" or a "student". All participants receive the same instructions and materials, however.** At the beginning of the experiment, your role is displayed on the screen. You are allowed to make notes during the whole experiment. You are informed about the remaining time of the single parts on the computer screen.

Part 1

In the first part of the experiment, **teachers** see a **presentation** which contains **28 slides** for a total of **20 minutes**. In this presentation, two topics - on 14 slides each - are explained. Teachers can control the presentation with the buttons at the bottom of the screen. Some additional materials belonging to the two topics are included with the instructions. Students receive the same additional materials. After the 20 minutes are over, teachers have 10 minutes to prepare for part 2 of the experiment. In total, part 1 therefore lasts 30 minutes.

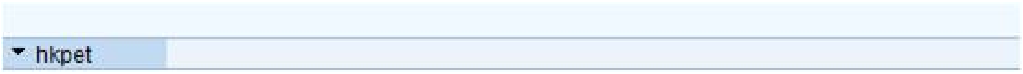
Students enter the laboratory during the course of the first part. They do **not** participate at part 1.

Part 2

Every teacher is randomly assigned one student. The computer automatically establishes a audio- and video-connection between a teacher and his student. Every teacher now has the chance to pass along the knowledge acquired in part 1 to his student. Picture and sound are transmitted only from the teacher. Each student can only confirm by clicking a button that he can see and hear the teacher at the beginning of the transmission. Before the transmission starts, teachers have one minute to adjust camera position and sound volume:

Camera position: Teachers can manually adjust the camera on the monitor in a way that they can see themselves in the middle of the right screen.

Sound volume: At the top of the video picture, teachers can as observe the volume level of their voice (dark blue area) when they speak into the microphone as seen on the following picture:



To adjust the volume level, a menu appears at the top of the screen when the mouse is moved there:

The indicator for the microphone (right) should be in a position that the volume level is clearly reacting when speaking normally.

Additionally, students can adjust the volume directly at the headset's cord.

In general, the content of the communication can be chosen freely by the teachers. They are however not permitted to give personal information about themselves. This includes for example name, age, address, study subject, or similar. Breaking this rule will lead to exclusion from the experiment and therefore no payoff.

After 10 minutes, the transmission stops automatically and part 3 starts.

Part 3

All participants complete a multiple-choice test which contains questions related to both topics from the presentation in part 1. The test consists of 30 questions overall, 15 for each topic. There are four possible answers for each question, **exactly one** of those is correct. There are no deductions for wrong answers. You have 40 seconds for every question, so part 3 lasts 20 minutes in total.

Payoff

The **teachers'** payoff is the sum of two components:

1. Teachers receive **EUR 0.75** for each question they have answered correctly themselves.
2. Teachers receive a fixed amount of EUR 4.50. [Only in treatment *Fix*]

2. Teachers receive **EUR 0.30** for each question their student has answered correctly. [only in treatment *Linear*]

2. Teachers receive **EUR 9**, if their student has answered at least 15 questions correctly. [only in treatment *Bonus*]

2. Three randomly chosen teachers form a group. The teacher in this group whose student answered the most questions correctly receives EUR 13.5. In the case of a tie, the **EUR 13.50** are divided equally among the respective teachers. [only in treatment *Tournament*]

Students receive EUR 0.75 for every question they have answered correctly themselves.

A.4 Appendix to Chapter 4

Derivation of the optimal values in the Minimal network

In the Minimal network the coordination problem is a small one. As a starting point, one optimal path in the Minimal network is depicted in Figure A.4.4. Red numbers indicate the step, black lines the allowed links and red lines the established connections. From this, the following reasoning shows the optimal behavior for all participants: In step 1 the decision of everyone but lotu is clear: Laga and lira both try to connect with lotu; lotu however can choose to either call laga or lira. A rational lotu would select randomly as there is no way to predict lelo's behavior in step 2. Lelo in step 2 does not know which of the two connections was established in step 1: lotu with laga or lotu with lira. Consequently lelo selects randomly. The decisions of the others are straight forward. Laga or lira (depending on step 1) respond to lelo, the others are off or do not matter at all. Now two scenarios can take place:

1. In step 2 a connection was established. In this case steps 3 and 4 are clear. The optimum is reached.
2. In step 2 no connection was established. In this case, lelo knows that the decision was unlucky and successfully tries the other option. Everything is thus postponed by one step as all the other players learn about the history during the run (e.g. by failed communication). The group reaches full information in 5 steps and with 6 connections.

In either case a rational group does not need more than 6 connections. For every subsequent run, however, it is important to distinguish whether the group reached the commonly known optimum in the run before. If they did, it is assumed that they are able to repeat it until the end of the experiment. If not, lotu in step 1 and lelo in step 2 again select randomly (as there is no better rationale). As a result, the group needs 6 connections in every round. How many steps are needed depends on the random decision described beforehand. Assuming that groups are able to stick to an optimal path if they have played it once, one can conclude that rational participants would need a total of 40.999 steps for 10 rounds in expectation. In the first round, one would expect 4.5 steps on average, in the last round 4.00098.

Instructions[115]

Welcome to the experiment!

You are participating in a study of decision-making behavior in the context of experimental economics. During the study you and the other participants will be asked to make decisions. You can earn money with this study. How much money you earn is dependent on the course of the experiment. You will receive detailed instructions about this in the following. All participants are paid in cash directly after the experiment one by one. To assure this, please remain seated after the experiment until your cabin number is called.

During the course of the experiment, no participant will receive information about the other participants' identities. All decisions are therefore made anonymously.

Should you have questions, please give a sign to alert one of the laboratory's employees. He will come to you and help you. **No communication with the other participants is allowed during the experiment; breaking this rule will lead to an immediate exclusion from the experiment.**

[115] These instructions are translated from the original, German instructions.

Instructions

The experiment consists of **10 rounds**. Before these rounds start, we would like to ask you to answer some comprehension questions. The experiment can only start when all participants have answered these questions correctly.

Initial situation

In the experiment, 5 randomly chosen participants form one group. This group stays the same for all 10 rounds. **Every group member** possesses **one piece of private information** at the beginning of each round. There are thus 5 different available pieces of information in each group. The group members form a network. During the course of a round the group members can establish connections with each other and exchange their information in that way. In the experiment, the possible connections which can be established in the network are shown on the screen. **Each circle represents a group member and each line represents a possible connection.** The names "lotu", "laga", "leje", "lira", and "lelo" identify the different positions in the network. Your own position in the network is marked red during the experiment.

Course of the experiment

Each round consists of several periods. At the beginning of each period every group member decides with which other group member he wants to establish a connection. It is also possible to establish no connection in a period. You choose the group member you want to establish a connection with by **clicking the corresponding circle** on the screen. The circle will then be colored blue. As soon as you confirm your choice, the connection attempt is started. If you do not want to establish a connection in this period, please click the corresponding button on the screen.

Only if **two group members choose each other** in one period, a **connection is actually established** and **all pieces of information** which both group members possess **are exchanged.** You are therefore not only passing along your own piece of private information but also - if present - the pieces of information of other participants if you have received them beforehand. In one period, each group member can only establish one connection at most. It is therefore also possible that no connection is established in a period.

A period ends when all participants have made their decision. It follows that all participants start each period at the same time. Starting with the second period, a table summarizing the earlier periods is displayed for every group member. This table also shows which pieces of information you currently possess and which group members have tried to establish a connection with you in the last periods.

A round ends when **all group members possess all 5 available pieces of information** or the deposit (see below) is depleted. A round does not yet end when one group member possesses all information. The number of periods per round is therefore not predetermined. At the end of each round, the table summarizing the periods is shown again, this time including, in addition, the payoff of the current round.

Participants' payoff

The **group's payoff** per round is calculated as follows: At the beginning of each round, every group has a **deposit of EUR 20**. In the course of a round, the following costs are subtracted:

- Each period costs EUR 0.30.
- Each established connection in the group costs EUR 0.8.

Trying to establish a connection without succeeding is costless. Periods however always cost EUR 0.30, even when no connections are established in them. The group's earnings are calculated by **subtracting all costs from the deposit** for one round. If the total costs from periods and established connections exceed the deposit, the round is aborted and the group's earnings for this round are 0.

The group's payoff is divided equally among the group members. The payoff of one group member for a round is therefore **one fifth of the group's payoff in this round.**

The total payoff for each participant is the sum of the payoffs in all 10 rounds. All members of a group therefore have the same payoff.

Figure A.4.1: Example of an optimal path in the Full network

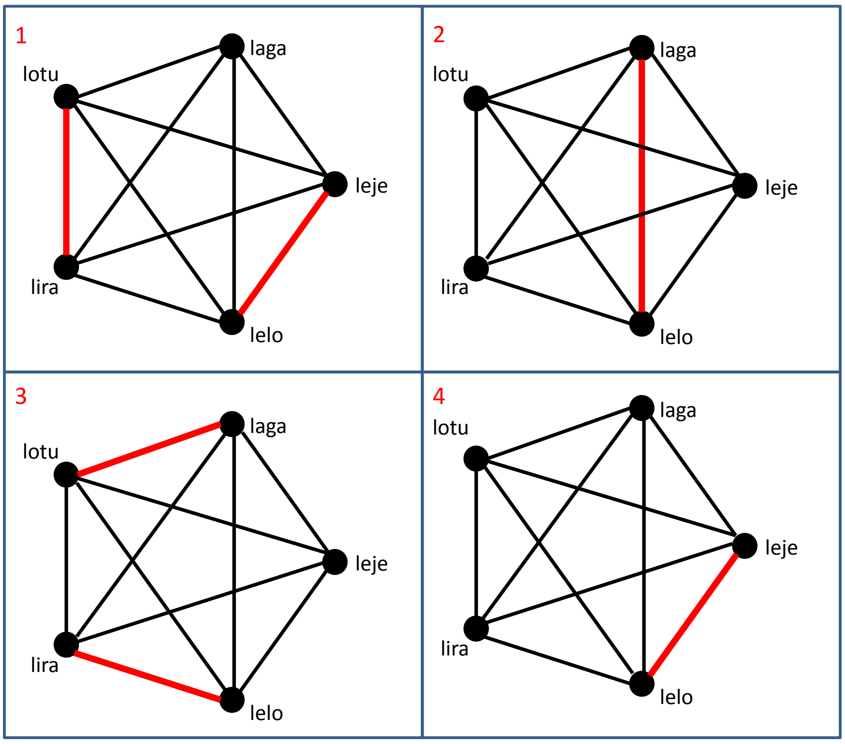


Figure A.4.2: Example of an optimal path in the Star network

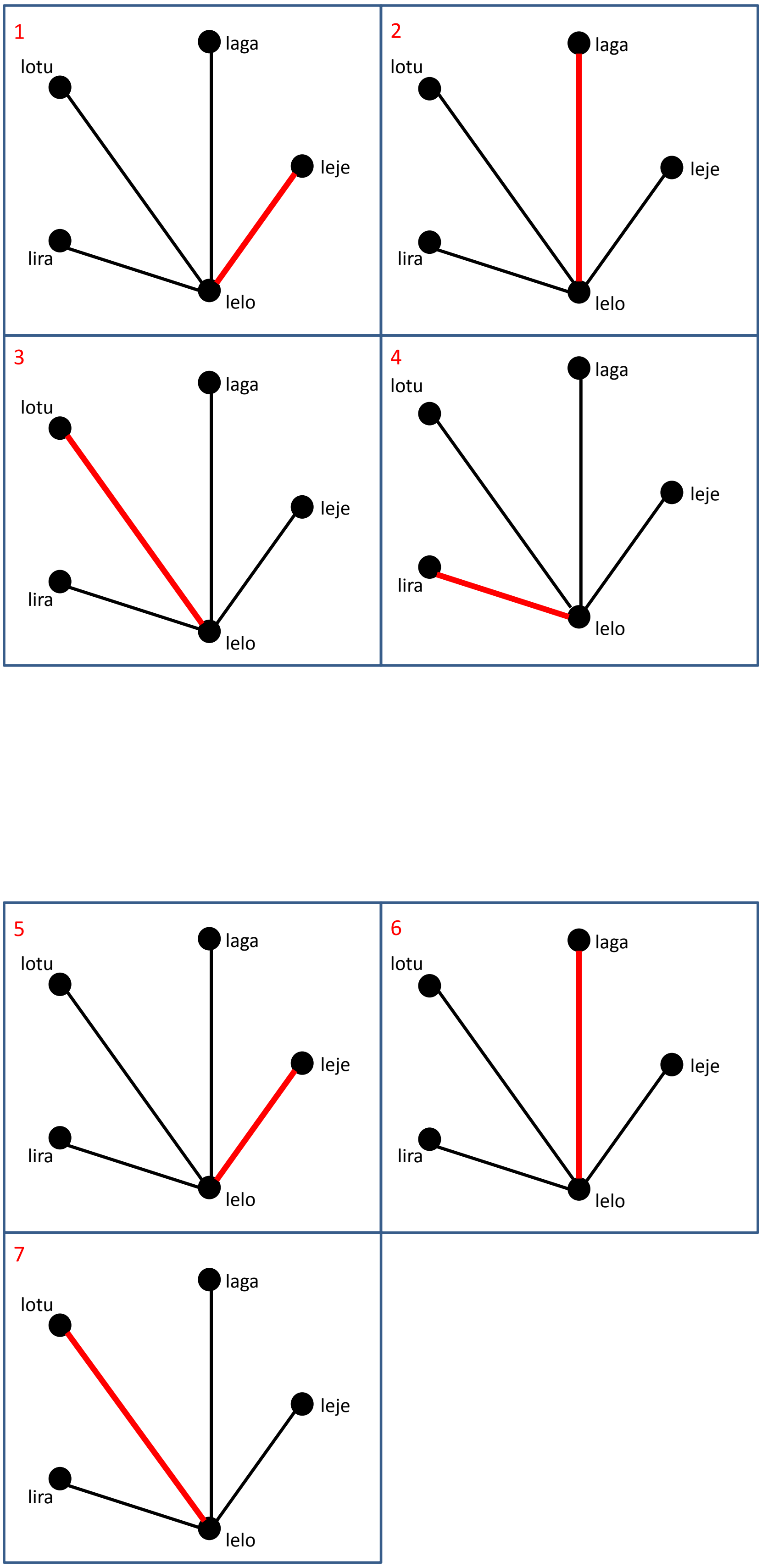


Figure A.4.3: Example of an optimal path in the Y network

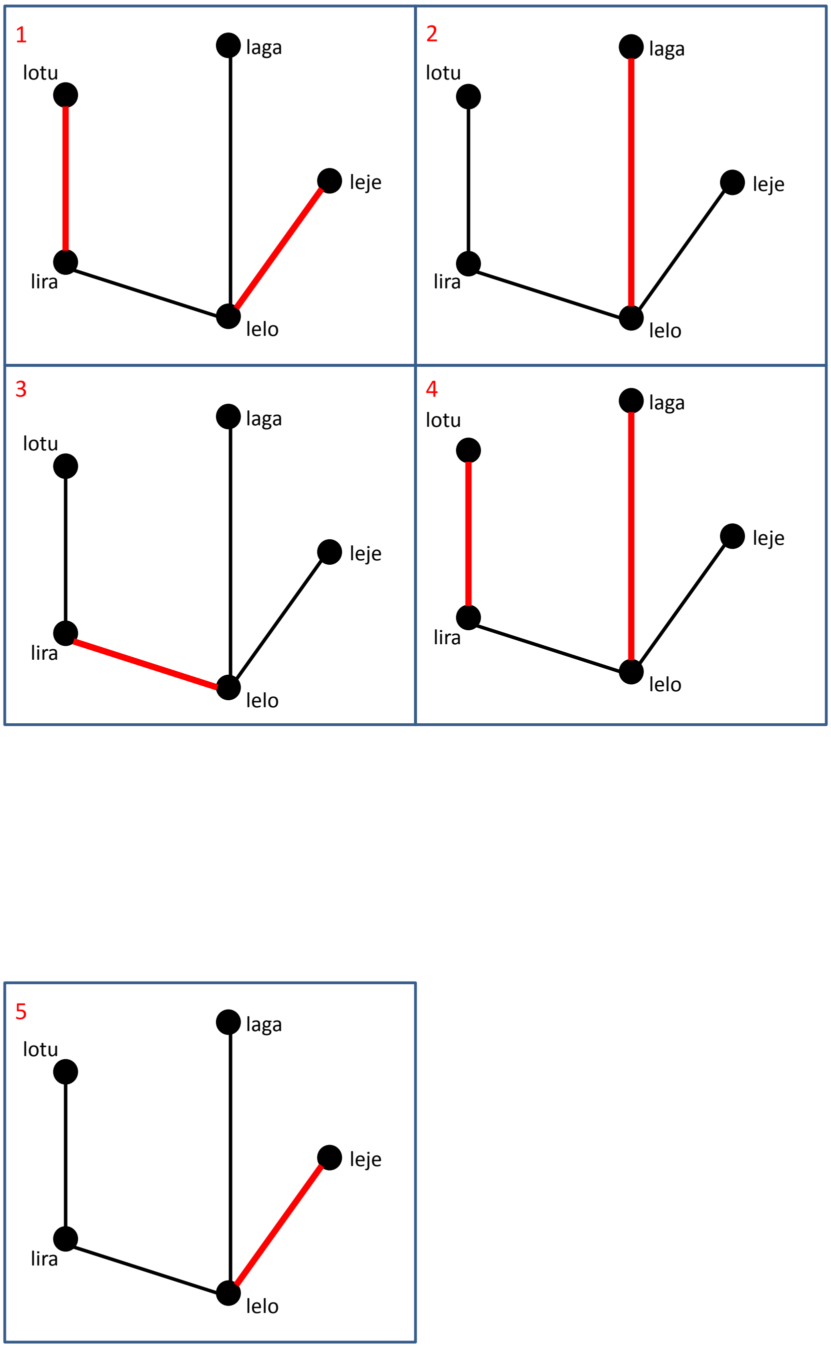


Figure A.4.4: Example of an optimal path in the Minimal network

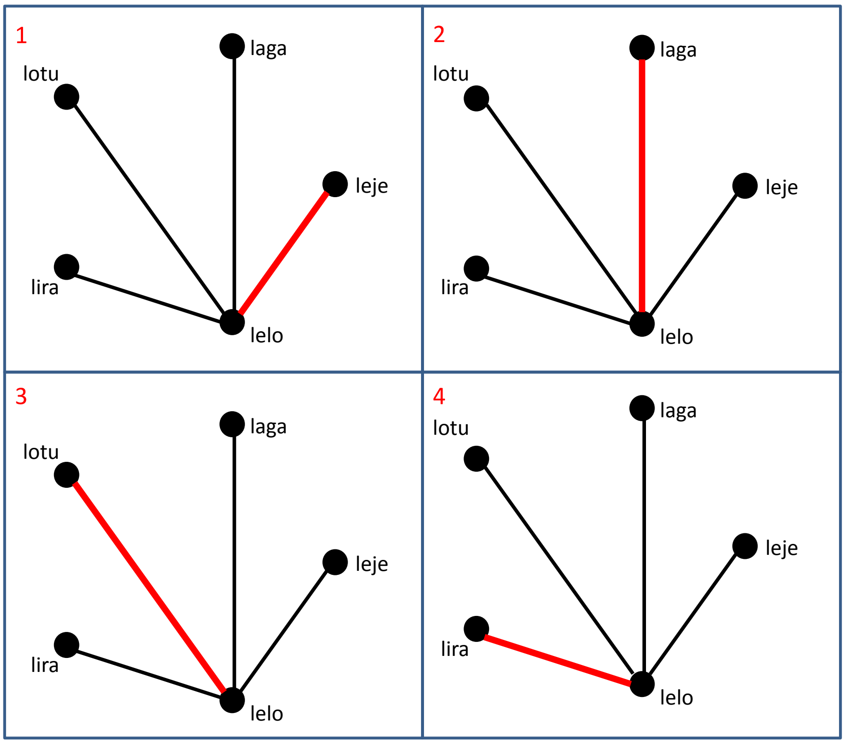


Figure A.4.5: Share of identically coded decisions

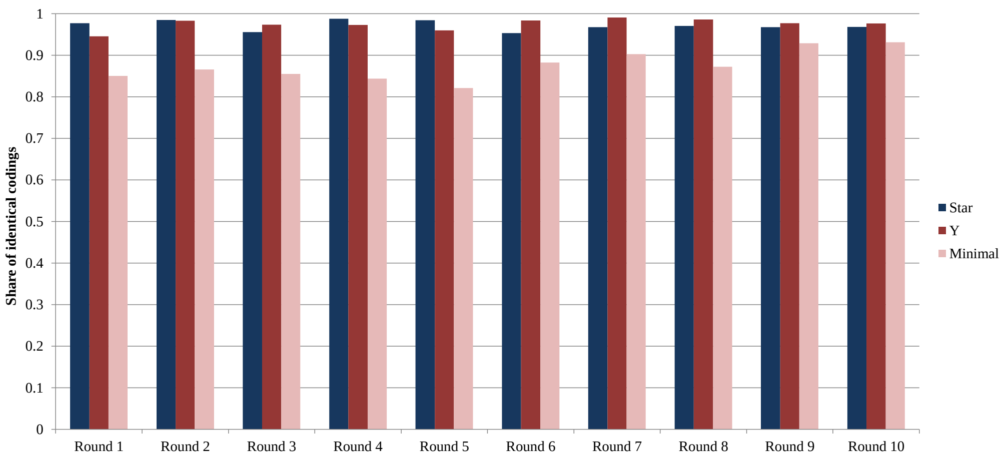


Table A.4.1: Questionnaire variables

Variable name	Question asked
Sex	Are you male or female?
Age	How old are you?
Study field	Which subject are you currently studying?
Length of study	How many semesters have you been studying your current subject?
Abi	What was your final grade point average in high school?
German	What was your last grade in high school in German?
Math	What was your last grade in high school in math?
Open start	Please think back to the beginning of the experiment. How did you decide to which group member you wanted to establish a connection in the very first period?
Closed diff	Did you proceed differently concerning establishing connections in the first one or two rounds than in the later rounds?
Open last	Please comment briefly on your decisions in the last round.
Closed happy	In how many rounds have you been content with the result?
Closed minperiods	In your opininon, how many periods are at least neccessary until every group member has all information?
Closed minconnections	In your opininon, how many connections are at least neccessary until every group member has all information?
Closed important1	Are all group members equally important for the exchange of information in your opinion?
Closed important2	If not: Which group member(s) is (are) the most important? (You can choose more than one)
Closed coop	Do you think that all group members worked together to exchange information?
Closed aspects 1 to 10	Which of the following aspects were important for your decisions with whom to establish a connection:
Closed aspects 1	Chance
Closed aspects 2	Experience from earlier rounds
Closed aspects 3	Connection attempts / order of connections from earlier periods
Closed aspects 4	Establish a connection with a group member whose information I am still missing
Closed aspects 5	Names (laga, leje, lotu, lira, lelo) of the other group members
Closed aspects 6	Geometrical position (“to the left of me”, “to the right of me”, etc.) of the other group members
Closed aspects 7	More / less important position of the other group members in the network
Closed aspects 8	Improving the result from the previous round
Closed aspects 9	Few alternatives to choose from
Closed aspects 10	Spread information, even without receiving new information myself

A.5 Appendix to Chapter 5

Instructions[116]

Welcome to the experiment!

You are participating at a study of decision making behavior in the context of experimental economics. During the study you and the other participants will be asked to make decisions. You can earn money with this study. How much money you earn is dependent on the course of the experiment. You will receive detailed instructions about this in the following. All participants are paid in cash directly after the experiment one by one. To assure this, please remain seated after the experiment until your cabin number is called.

Should you have any questions before the start of the experiment, please ask an employee of the laboratory. He will come to your place and help you. **Any communication with the other participants during the experiment is only allowed when explicitly prompted; breaking this rule will lead to an immediate exclusion from the experiment.**

[116]These instructions are translated from the original, German instructions.

Instructions

The experiment consists of **thirty rounds**. In each of these rounds you have to decide how to split your available money between **saving** and **spending**. The current round is displayed at the top of the screen. Your money is denoted in the unit **experimental currency (EW)**.

Available money

Your **available money** consists of two parts: Your **saving**s from earlier rounds and your **current income**. It is denoted in EW.

Available money = Savings + current income

Current income

In each round, you receive a **current income**. This current income also consists of two parts: a **fixed income** and a **variable adjustment factor**.

Current income = Fixed income x variable adjustment factor

The fixed income is 100 EW in the first round and increases by 5% in each round afterwards. In the second round your fixed income is therefore 105 EW, in the third round 110.25 EW etc.

Variable adjustment factor

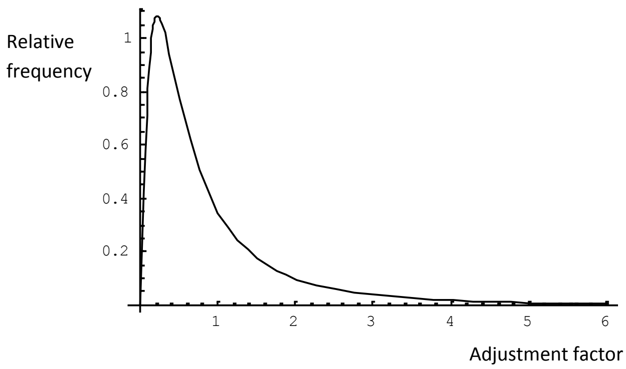
The adjustment factor is variable, because in each round there is a chance to become **unemployed**. A low adjustment factor - and the resulting low current income - therefore represents high unemployment.

The adjustment factor is variable, because the **overall economic situation** is taken into account in each round. A low adjustment factor - and the resulting low current income - therefore represents a bad economic situation.

The adjustment factor is variable, because in each round the **tax policy** can be changed. A low adjustment factor - and the resulting low current income - therefore represents tax policy which puts high taxes on income.

The adjustment factor is determined using a probability distribution, therefore it is determined **randomly**. The adjustment factors are determined independently in each round. A certain round's adjustment factor therefore has no influence on the next round's adjustment factor.

The probability distribution is represented by the following density function:



The x-axis shows all possible adjustment factors; the y-axis shows the relative frequency of each value. The adjustment factor is therefore always larger than 0, but will be smaller that 1 in most cases. In about 10% of all cases the adjustment factor will be smaller than 0.168. In about 10% of all cases it will be larger than 2.185. In ca. 80% of all cases the adjustment factor will be between 0.168 and 2.185.

The following table shows three possible sequences of adjustment factors, which have been generated with this probability distribution. These possible courses are not used in the experiment and only serve as examples.

	Sequence A	**Sequence B**	**Sequence C**
Round	Adjustment factor	Adjustment factor	Adjustment factor
1	1.364	0.845	0.624
2	0.461	2.464	2.660
3	0.498	0.403	2.643
4	0.223	0.199	1.298
5	0.323	0.413	0.840
6	0.108	0.296	0.389
7	0.283	0.199	0.530
8	0.588	0.926	2.592
9	4.793	1.989	0.599
10	0.780	1.601	1.246
11	2.721	0.230	0.674
12	0.334	1.270	0.159
13	2.203	0.715	1.586
14	1.363	0.404	0.129
15	0.289	0.100	0.471
16	0.194	0.170	0.309
17	0.369	0.426	0.364
18	1.296	0.604	0.703
19	0.256	0.248	1.120
20	0.308	1.033	0.219
21	0.767	1.441	0.780
22	0.671	0.910	0.049
23	0.578	0.198	0.486
24	0.956	1.665	0.446
25	2.000	1.636	0.265
26	1.782	0.174	0.549
27	0.140	0.482	0.276
28	0.384	0.342	0.406
29	0.087	0.929	0.457
30	1.692	1.625	0.367

The expected value of the adjustment factor is 1. When many adjustment factors are determined randomly, the average of their values will be 1.

The expected value of the adjustment factor is 1. When many adjustment factors are determined randomly, the average of their values will be 1. In addition, in about 10% of all cases the adjustment factor will be smaller than 0.168. In about 10% of all cases it will be larger than 2.185. In ca. 80% of all cases the adjustment factor will be between 0.168 and 2.185.

Your estimation of the adjustment factor

Before your decision about saving and spending you have to estimate in every round, which adjustment factor will be **determined in the current round**. Please enter your estimation in the field "Estimation" on the screen.

Before your decision you are informed about the estimations about the adjustment factor which have been made by four other participants. These four participants are assigned randomly to you and stay the same for all 30 rounds.

Information after each round

In each round, a table shows savings, fixed income, adjustment factor, and the resulting available money. The values of the past rounds are also displayed in the table.

Your decision

Please enter your **decision** how much of your available money you want to spend in the current round (and with it indirectly how much you want to save) in the field "Your decision".

The amount of EW which you **spend** in each round is converted to **Eurocent**.

The lifestyle index

For **testing purposes**, you can check how many Eurocents you receive for a certain amount of EW with the Euro calculator on the screen.

The conversion from EW to Eurocents is influenced by your **lifestyle index**. The **higher** the lifestyle index, **the less** Eurocent you receive for a certain amount of EW you spend.

The appendix of these instructions contains a **Eurocent conversion table (Table 1)** which shows how spending is converted to Eurocent. The following table is an excerpt of this conversion table:

Spending (in EW)	Lifestyle index 10	20	50	100	150	200	250	300
5	-391.57	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00
10	20.48	-290.27	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00
20	185.03	87.76	-257.20	-400.00	-400.00	-400.00	-400.00	-400.00
40	238.81	211.32	113.83	-75.81	-286.26	-400.00	-400.00	-400.00
60	250.17	237.42	192.21	104.26	6.65	-97.80	-207.67	-322.04
80	254.35	247.02	221.03	170.48	114.37	54.33	-8.82	-74.56
100	256.34	251.59	234.73	201.95	165.57	126.64	85.69	43.06
120	257.43	254.10	242.30	219.33	193.84	166.57	137.88	108.01
140	258.10	255.64	246.91	229.93	211.09	190.92	169.71	147.63
160	258.54	256.65	249.93	236.87	222.37	206.86	190.55	173.56
180	258.84	257.34	252.02	241.66	230.16	217.86	204.92	191.45
200	259.06	257.84	253.51	245.10	235.76	225.76	215.25	204.31
220	259.22	258.21	254.63	247.65	239.92	231.64	222.93	213.86
240	259.34	258.49	255.48	249.61	243.09	236.12	228.79	221.15
260	259.44	258.71	256.14	251.13	245.57	239.62	233.36	226.84

For example, if you decide to spend 100 EW in a round, you will receive 256.34 Eurocent if your lifestyle-index is 10, but only 201.95 Eurocent if your lifestyle-index is 100.

Development of the lifestyle index

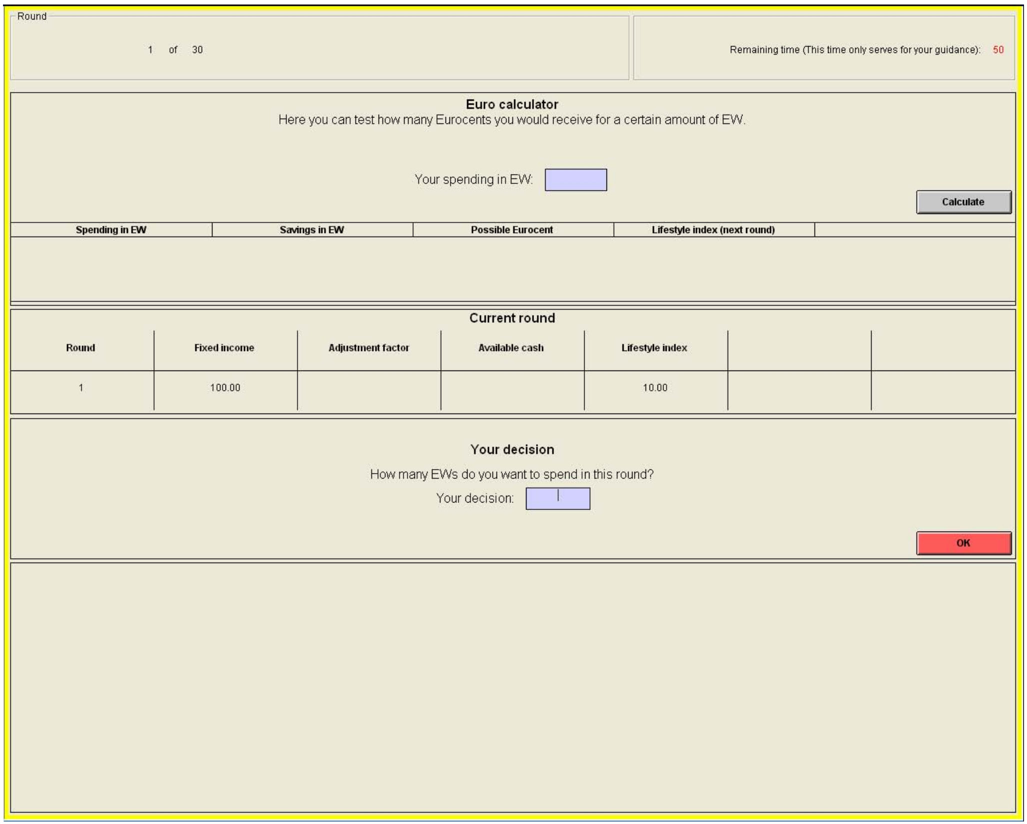
In each round, the lifestyle index is calculated as follows:

Lifestyle index (current round) = 0.7 x lifestyle index (last round) + Spending (last round)

Therefore, the **more** you spend in each round, the **higher** the lifestyle index gets. At the end of these instructions you find a **lifestyle table (Table 2)** which shows the development of the lifestyle index for a number of spending amounts. The following table is an excerpt of this table:

Lifestyle index, current round

Spending, current round (in EW)	10	20	50	100	150	200	250
10	17	24	45	80	115	150	185
20	27	34	55	90	125	160	195
40	47	54	75	110	145	180	215
60	67	74	95	130	165	200	235
80	87	94	115	150	185	220	255
100	107	114	135	170	205	240	275
120	127	134	155	190	225	260	295
140	147	154	175	210	245	280	315
160	167	174	195	230	265	300	335
180	187	194	215	250	285	320	355
200	207	214	235	270	305	340	375
220	227	234	255	290	325	360	395
240	247	254	275	310	345	380	415
260	267	274	295	330	365	400	435
280	287	294	315	350	385	420	455
300	307	314	335	370	405	440	475
320	327	334	355	390	425	460	495



If for example you choose to spend an amount of 60 EW with a lifestyle index of 10, your lifestyle index in the **next round** will be 67.

The lifestyle index in the **first round is 10**.

Course of the rounds

As already described, you have to decide in each of the 30 rounds how to split your available money between spending and saving. Once you have decided upon a split and left the round, you **cannot change** this decision any more. I.e. you cannot return to past rounds. The following picture shows the your decision screen in the experiment.

Your payoff

After the 30 rounds, **all Eurocents** which you received through your spending are **added up**. You are payed this **sum** in cash at the end of the experiment.

In addition, you receive a **show-up fee of EUR 5** for your participation at the experiment.

Comprehension questions

Before the experiment starts, you answer some comprehension questions on the screen. The experiment will only start, when you have answered all questions correctly. These questions do not influence your payoff. **If you have questions regarding the instructions, please raise your hand.** An employee of the laboratory will come to you and answer your questions.

Eurocent conversion table (Table 1)

Lifestyle index

Spending (in EW)	10	20	50	100	150	200	250	300	350	400	500	600	700	…
5	-391.57	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	
10	20.48	-290.27	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	
20	185.03	87.76	-257.20	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	
40	238.81	211.32	113.83	-75.81	-286.26	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	-400.00	
60	250.17	237.42	192.21	104.26	6.65	-97.80	-207.67	-322.04	-400.00	-400.00	-400.00	-400.00	-400.00	
80	254.35	247.02	221.03	170.48	114.37	54.33	-8.82	-74.56	-142.55	-212.50	-357.59	-400.00	-400.00	
100	256.34	251.59	234.73	201.95	165.57	126.64	85.69	43.06	-1.03	-46.39	-140.47	-238.41	-339.68	
120	257.43	254.10	242.30	219.33	193.84	166.57	137.88	108.01	77.13	45.35	-20.56	-89.17	-160.12	
140	258.10	255.64	246.91	229.93	211.09	190.92	169.71	147.63	124.80	101.30	52.58	1.85	-50.61	
160	258.54	256.65	249.93	236.87	222.37	206.86	190.55	173.56	156.00	137.92	100.44	61.41	21.06	
180	258.84	257.34	252.02	241.66	230.16	217.86	204.92	191.45	177.52	163.19	133.46	102.51	70.51	
200	259.06	257.84	253.51	245.10	235.76	225.76	215.25	204.31	192.99	181.35	157.20	132.06	106.06	
220	259.22	258.21	254.63	247.65	239.92	231.64	222.93	213.86	204.49	194.84	174.83	154.00	132.47	
240	259.34	258.49	255.48	249.61	243.09	236.12	228.79	221.15	213.26	205.14	188.29	170.75	152.62	…
260	259.44	258.71	256.14	251.13	245.57	239.62	233.36	226.84	220.11	213.17	198.79	183.83	168.35	
280	259.52	258.89	256.67	252.34	247.54	242.40	237.00	231.37	225.55	219.56	207.15	194.22	180.86	
300	259.58	259.03	257.09	253.32	249.13	244.65	239.93	235.03	229.95	224.73	213.90	202.63	190.97	
325	259.64	259.17	257.52	254.30	250.73	246.90	242.88	238.69	234.36	229.91	220.67	211.05	201.10	
350	259.69	259.29	257.86	255.08	251.99	248.69	245.22	241.61	237.87	234.02	226.05	217.74	209.15	
375	259.73	259.38	258.13	255.71	253.02	250.14	247.11	243.96	240.70	237.35	230.39	223.15	215.66	
400	259.76	259.45	258.36	256.22	253.86	251.33	248.66	245.89	243.02	240.07	233.95	227.58	221.00	
425	259.79	259.51	258.54	256.65	254.56	252.31	249.95	247.49	244.95	242.33	236.91	231.26	225.42	
450	259.81	259.57	258.70	257.01	255.14	253.14	251.03	248.83	246.57	244.23	239.39	234.35	229.14	
475	259.83	259.61	258.83	257.32	255.64	253.84	251.94	249.97	247.94	245.84	241.49	236.96	232.28	
500	259.85	259.65	258.95	257.58	256.06	254.43	252.72	250.95	249.11	247.21	243.29	239.20	234.97	
550	259.87	259.71	259.13	258.00	256.74	255.40	253.98	252.51	250.99	249.42	246.17	242.79	239.29	
600	259.89	259.76	259.27	258.31	257.26	256.13	254.94	253.70	252.42	251.10	248.37	245.53	242.59	
650	259.91	259.79	259.37	258.56	257.66	256.70	255.68	254.63	253.54	252.41	250.09	247.66	245.15	
700	259.92	259.82	259.46	258.76	257.98	257.15	256.28	255.37	254.42	253.46	251.45	249.35	247.20	
⋮							⋮							⋱

<u>Lifestyle table (Table 2)</u>

Lifestyle index, current round

Spending amount, current round (in EW)

	10	**20**	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**500**	**600**	**700**	**800**	**900**	**…**
10	17	24	45	80	115	150	185	220	255	290	360	430	500	570	640	
20	27	34	55	90	125	160	195	230	265	300	370	440	510	580	650	
40	47	54	75	110	145	180	215	250	285	320	390	460	530	600	670	
60	67	74	95	130	165	200	235	270	305	340	410	480	550	620	690	
80	87	94	115	150	185	220	255	290	325	360	430	500	570	640	710	
100	107	114	135	170	205	240	275	310	345	380	450	520	590	660	730	
120	127	134	155	190	225	260	295	330	365	400	470	540	610	680	750	
140	147	154	175	210	245	280	315	350	385	420	490	560	630	700	770	
160	167	174	195	230	265	300	335	370	405	440	510	580	650	720	790	
180	187	194	215	250	285	320	355	390	425	460	530	600	670	740	810	
200	207	214	235	270	305	340	375	410	445	480	550	620	690	760	830	
220	227	234	255	290	325	360	395	430	465	500	570	640	710	780	850	
240	247	254	275	310	345	380	415	450	485	520	590	660	730	800	870	
260	267	274	295	330	365	400	435	470	505	540	610	680	750	820	890	
280	287	294	315	350	385	420	455	490	525	560	630	700	770	840	910	…
300	307	314	335	370	405	440	475	510	545	580	650	720	790	860	930	
320	327	334	355	390	425	460	495	530	565	600	670	740	810	880	950	
340	347	354	375	410	445	480	515	550	585	620	690	760	830	900	970	
360	367	374	395	430	465	500	535	570	605	640	710	780	850	920	990	
380	387	394	415	450	485	520	555	590	625	660	730	800	870	940	1010	
400	407	414	435	470	505	540	575	610	645	680	750	820	890	960	1030	
420	427	434	455	490	525	560	595	630	665	700	770	840	910	980	1050	
440	447	454	475	510	545	580	615	650	685	720	790	860	930	1000	1070	
460	467	474	495	530	565	600	635	670	705	740	810	880	950	1020	1090	
480	487	494	515	550	585	620	655	690	725	760	830	900	970	1040	1110	
500	507	514	535	570	605	640	675	710	745	780	850	920	990	1060	1130	
520	527	534	555	590	625	660	695	730	765	800	870	940	1010	1080	1150	
540	547	554	575	610	645	680	715	750	785	820	890	960	1030	1100	1170	
560	567	574	595	630	665	700	735	770	805	840	910	980	1050	1120	1190	
580	587	594	615	650	685	720	755	790	825	860	930	1000	1070	1140	1210	
600	607	614	635	670	705	740	775	810	845	880	950	1020	1090	1160	1230	
620	627	634	655	690	725	760	795	830	865	900	970	1040	1110	1180	1250	
⋮								⋮								⋱